

Maximum Entropy for Gravitational Wave Data Analysis: Inferring the Physical Parameters of Core-Collapse Supernovae

T. Z. Summerscales¹

Center for Gravitational Wave Physics, Penn State University, University Park, PA 16802

tzs@gravity.psu.edu

Adam Burrows

Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721

burrows@zenith.as.arizona.edu

Lee Samuel Finn

Center for Gravitational Wave Physics, Penn State University, University Park, PA 16802

lsfinn@psu.edu

and

Christian D. Ott

Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721

cott@as.arizona.edu

ABSTRACT

The gravitational wave signal arising from the collapsing iron core of a Type II supernova progenitor star carries with it the imprint of the progenitor's mass, rotation rate, degree of differential rotation, and the bounce depth. Here, we show how to infer the gravitational radiation waveform of a core collapse event from noisy observations in a network of two or more LIGO-like gravitational wave detectors and, from the recovered signal, constrain these source properties. Using these techniques, predictions from recent core collapse modeling efforts, and the LIGO performance during its S4 science run, we also show that gravitational wave

¹Now at Department of Physics, Andrews University, Berrien Springs, MI, 49104

observations by LIGO might have been sufficient to provide reasonable estimates of the progenitor mass, angular momentum and differential angular momentum, and depth of the core at bounce, for a rotating core collapse event at a distance of a few kpc.

Subject headings: gravitational waves — methods: data analysis — supernovae: general

1. INTRODUCTION

The gravitational waves that we expect to observe in large detectors such as LIGO (Waldman & The LIGO Scientific Collaboration 2006; Mandic 2006), Virgo (Acernese et al. 2006), GEO600 (Lück et al. 2006), TAMA300 (Takahashi & the TAMA Collaboration 2004) and LISA (Robertson 2000; Rüdiger 2004), reflect the coherent evolution of the most compact part of the source mass distribution. From the observed waves we have the potential to infer the factors that govern that evolution. For example, the evolution of the collapsing stellar core in a type II supernova is determined in part by the progenitor mass density and angular momentum distributions in the inner core. None of these properties of the progenitor can be directly determined from the electromagnetic radiation we observe when the shock emerges from the stellar envelope. In addition, the forces that govern the evolution of the core depend upon many things that are unknown, or poorly modeled, today (for example, the matter equation of state, the role played by neutrinos and neutrino radiation transport, general relativity, convection and non-axisymmetry). The gravitational waves emitted during the collapse phase and its aftermath carry the signature of all these preconditions and the dynamics that govern both the collapse and subsequent rebound. Reading that signature requires inferring the waveform from the noisy detector observations. Here, we develop a new method for inferring the gravitational radiation waveform from the noisy data from two or more detectors, based on Jaynes’ principle of maximum entropy (Jaynes 1957b,a), demonstrate its effectiveness when applied to the discovery of signals arising from simulated rotating iron core collapse buried in simulated detector noise, and show how the inferred waveform can be used, in principle, to gain insight into the properties of the source.

The first problem addressed in this study is the problem of inferring the gravitational wave signal from the data produced by the detectors, also referred to as the deconvolution or inverse problem. Inverse problems generally have been long-recognized as a problem to be approached with great care (see Evans & Stark 2002 for a recent review). The detector response, which relates the incident gravitational wave to the signal observed in the detector, is generally an ill-conditioned function; additionally, the presence of additive noise generally

confuses the observation. Together, the noise and the ill-conditioned detector response generally thwart naïve attempts at signal deconvolution. Early attempts at gravitational wave signal deconvolution explored a least-squares or maximum likelihood approach to deconvolution: e.g., Gürsel & Tinto (1989) developed a procedure for inferring a plane wave signal incident on a network of detectors with a frequency-independent response, in which case the problem response is well-conditioned. These techniques tend to over-fit the observations; additionally, realistic detectors have frequency responses that are ill-conditioned and this complicates a least-squares or maximum likelihood approach to deconvolution. More recent efforts have explored regularized methods for inferring the incident gravitational radiation waveform (Rakhmanov 2006). Here, we use the *maximum entropy* principle to regularize the deconvolution problem, developing an application to the gravitational wave inference problem that is applicable in the general case of a frequency dependent and ill-conditioned detector response and avoids over-fitting the observations in the presence of additive detector noise. Going beyond the problem of inferring the incident wave, we use the inferred waveform to explore how, and how well gravitational wave observations of this kind can be used to learn about the source.

Maximum entropy approaches to deconvolution have a long heritage in astronomical image reconstruction (Ponsonby 1973; Gull & Daniell 1978, 1979; Skilling & Bryan 1984; Steenstrup 1985; Shevgaonkar 1987; Nityananda & Narayan 1982, 1983; Narayan & Nityananda 1984, 1986; Pantin & Starck 1996; Starck & Pantin 1996, 1997; Barreiro et al. 2001; Maisinger et al. 2004), where they have been used in all wavebands. Recent examples of the use of maximum entropy based methods include the reconstruction of the cosmic microwave background (Maisinger et al. 1997; Vielva et al. 2001) including maps based on data from WMAP (Bennett et al. 2003) and COBE (Jones et al. 1998, 1999; Barreiro et al. 2004).¹ Our work applies that heritage to the simpler problem of reconstructing the time-dependent plane wave signal incident on a network of gravitational wave detectors from their time series response.

Once we have inferred the incident signal, we are faced with a second problem of inference: identifying the properties of the source from the signal. One way of associating gravitational waveforms with supernova properties is to compare the inferred waveform with models arising from simulations that explore the signal dependency over a broad range of physical parameters. We should expect the inferred waveform to have the most in common with the simulated waveforms arising from models whose character is most similar to

¹In all these applications the Principle of Maximum Entropy actually plays a relatively small role: much more important is Bayes Law and the understanding that probability can represent degree of belief or certainty in the state of a system; nevertheless, the term Maximum Entropy has come to refer to all these methods of deconvolution and we continue that tradition.

the actual source. Here, we compare, using the cross-correlation, inferred waveforms with simulated waveforms arising from different rotating core collapse models. We assume that models whose signal shows the greatest correlation are the ones most likely to be similar to the source.²

Section 2 describes in detail the development of the inference Bayesian inference method that we have developed to identify the time-dependent gravitational wave signal incident on a network of detectors, ending with a demonstration of the method applied to simulated observations made at an idealized two-detector network. Section 3 reviews the expected gravitational-wave emission processes in core-collapse supernovae and describes the recently produced catalog of rotating core-collapse waveforms by Ott et al. (2004) and the physics that went into the considered models. We use these simulated core-collapse signals to characterize how well our procedure for inferring the incident signal and characterizing the source works. Finally, in § 4 we summarize our conclusions and directions for further study.

One important goal with this paper is to connect the two communities of gravitational wave experimentalists and supernova modelers in a way that has not been done in the past. For this reason we have taken particular care to make the discussion of § 2 pedagogical in nature. Future gravitational wave models can be put through the pipeline established with this paper so that we can obtain more credible estimates of what might be possible with either initial or advanced LIGO, Virgo, GEO600, or TAMA300, or with any combination of these, for any theoretical model. This has simply never been done for any collapse models, and is a novelty of this paper that we hope will stimulate further interactions between astrophysicists doing supernova simulations and the gravitational wave detection and data-analysis communities.

2. A Bayesian Approach to Deconvolution

2.1. Introduction

The first problem faced in the analysis and interpretation of astronomical data is the identification of a signal in noisy observations. When the signal being sought takes a known

²Note the distinction between correlation and the more involved matched filtering (Finn 1992; Finn & Chernoff 1993). Matched filtering is a useful analysis technique when the functional form of the signal being sought is known precisely. Here, we presume that we have only a qualitative model of the signal dependence on the physical parameters of interest. In that case, using the full apparatus of matched filtering could very well lead us to reject real signals because the match to the model is only qualitative.

form then special techniques, tuned to the characteristics of the signal, may be used; however, the general problem remains the same: the first requirement is to identify the characteristics of the radiation \mathbf{h} incident on the detector(s) that gives rise to the observational data \mathbf{d} . The data may be from a radio antenna, a spectrograph, or a network of gravitational wave detectors; our goal may be to map of the radio emission from a distant AGN, the emission spectra of an accretion disk, or the burst of gravitational waves associated with a core-collapse supernova. In the case of gravitational wave data, which is our principle interest in this paper, the data \mathbf{d} are a set of time series corresponding to the sampled output of each detector, one time series from each detector in our network; the radiation \mathbf{h} is the time dependent gravitational wave strain incident as a plane wave on the detector(s), which may be characterized (for example) by the direction of wave propagation and the time dependent strain amplitudes $h_+(t)$ and $h_\times(t)$ in the two polarization states as measured at the Earth’s barycenter. Restricting attention to linear detectors (which is almost always the case and thus hardly a restriction at all) the observations in each case are a linear superposition of a confounding noise and the incident radiation as filtered by the detector response function \mathbf{R} :

$$\mathbf{d} = \mathbf{R}\mathbf{h} + \mathbf{n} \quad (1)$$

where $\mathbf{R}\mathbf{h}$ is the convolution of the detector network response with the incident radiation and \mathbf{n} is the noise contribution to the observation \mathbf{d} . We generally know \mathbf{R} and the statistical properties \mathcal{N} of the noise \mathbf{n} ; with these our aim is to determine \mathbf{h} from \mathbf{d} .

In the absence of noise the solution to this problem appears, at first blush, to be straightforward: i.e., we solve for \mathbf{h} the system of equations 1 with $\mathbf{n} = 0$. Problems arise when \mathbf{R} is not invertible, either because we have too much or two little data (i.e., the problem is over- or under-determined), or because the detector(s) is (are) insensitive to some aspects of the signal \mathbf{h} (e.g., the detector cannot distinguish between polarization states of \mathbf{h} , or is insensitive to signal energy outside of some band). In these cases we are forced to deal with uncertainty: if the data \mathbf{d} over-determine \mathbf{h} we must be prepared to resolve contradictions within the data; if \mathbf{d} under-determine \mathbf{h} any conclusions we reach must be tempered by our incomplete knowledge of \mathbf{h} .

These problems are only compounded when we consider the real case of noisy data. In real problems the response function \mathbf{R} is always ill-conditioned: i.e., the response is not invertible (the problem is over-determined), or not unique (the problem is under-determined), or the solution \mathbf{h} to an equation of the form $\mathbf{R}\mathbf{h} = \mathbf{d}$ is sensitive to small perturbations in \mathbf{d} . Sensitivity of \mathbf{h} to perturbations in \mathbf{d} presents a new twist: even when the noise contribution to the observation \mathbf{d} is numerically small, it is dangerous to assume that $\mathbf{R}^{-1}\mathbf{d}$ is close to \mathbf{h} . Finally, even in those cases where the response is invertible and well-conditioned, applying the inverse to \mathbf{d} does not distinguish between the signal \mathbf{h} and a “signal equivalent

noise” $\mathbf{R}^{-1}\mathbf{n}$.

Framing the problem of deconvolution in the language of Bayesian inference provides guidance on how to proceed in determining \mathbf{h} . In a Bayesian approach to the problem of inference we seek not \mathbf{h} , but a probability distribution

$$f(\mathbf{h}'|\mathbf{d}, \mathbf{R}, \mathcal{N}, \mathcal{I}) = \begin{pmatrix} \text{probability density that } \mathbf{h}' \text{ is incident on the detector(s)} \\ \text{given data } \mathbf{d}, \text{ response } \mathbf{R}, \text{ noise characterization} \\ \mathcal{N}, \text{ and other unenumerated assumptions } \mathcal{I}. \end{pmatrix}. \quad (2)$$

The probability density f , if we can find it, fully describes our legitimate knowledge — including uncertainty — regarding the waveform and its properties: e.g., we can compute from it our expectation of the gravitational wave power incident on the detector, etc. It also provides a good point estimate of the incident radiation: i.e., the \mathbf{h}' that maximizes $f(\mathbf{h}'|\mathbf{d}, \mathbf{R}, \mathcal{N}, \mathcal{I})$. We can report f or, as is more generally the case, report some summary of the distribution f : e.g., the \mathbf{h}' that maximizes f and some suitably defined “error bars”, which summarize the degree of our uncertainty.

Following Bayes Law, the probability distribution f may be “factored” into the product of three other distributions:

$$f(\mathbf{h}'|\mathbf{d}, \mathbf{R}, \mathcal{N}, \mathcal{I}) = \frac{g(\mathbf{d}|\mathbf{h}', \mathbf{R}, \mathcal{N}, \mathcal{I})q(\mathbf{h}'|\mathcal{I})}{v(\mathbf{d}|\mathbf{R}, \mathcal{N}, \mathcal{I})}, \quad (3a)$$

where

$$g(\mathbf{d}|\mathbf{h}', \mathbf{R}, \mathcal{N}, \mathcal{I}) = (\text{likelihood function for } \mathbf{d} \text{ given } \mathbf{h}') \quad (3b)$$

$$q(\mathbf{h}'|\mathcal{I}) = (\text{a priori} \text{ expectations regarding the incident wave.}) \quad (3c)$$

$$v(\mathbf{d}|\mathbf{R}, \mathcal{N}, \mathcal{I}) = \begin{pmatrix} \text{normalization constant equal to the probability} \\ \text{of observing } \mathbf{d} \text{ given response } \mathbf{R} \text{ and noise } \mathcal{N} \end{pmatrix} \quad (3d)$$

The probability v is independent of \mathbf{h} . We can ignore it when our only goal is to find the \mathbf{h} that maximizes f ; however, as we will see below, v plays a vitally important role in helping us choose our prior q . In the following subsections we describe how to compute f by finding, in turn, of the probabilities g , q and v .

2.2. The Likelihood Function g

Focus attention first on the likelihood function $g(\mathbf{d}|\mathbf{h}', \mathbf{R}, \mathcal{N}, \mathcal{I})$. If the signal content of \mathbf{d} corresponds to incident gravitational radiation \mathbf{h} , then $\mathbf{d} - \mathbf{R}\mathbf{h}$ is just detector noise.

Correspondingly, the likelihood function $g(\mathbf{d}|\mathbf{h}, \mathbf{R}, \mathcal{N}, \mathcal{I})$ is just the probability that $\mathbf{d} - \mathbf{R}\mathbf{h}$ is noise. How do we evaluate this probability?

Gravitational wave detectors, interferometric or acoustic, measure a continuous quantity: i.e., they are not particle detectors that count discrete events (e.g., photons). The statistical properties of the noise in these detectors is characterized by its power- and cross-spectral density (Abbott et al. 2004a; Lazzarini 2002, 2003, 2004, 2005) or, equivalently, its mean and covariance (Kittel 1958).³ The Principle of Maximum Entropy was introduced by Jaynes (1957b,a) as a means of identifying probability distributions whose “information content” is, in a well-defined and relevant way, consistent with this kind of a priori information but otherwise incorporates no other assumptions. It was first articulated in the context of statistical thermodynamics, where it provided a foundation for understanding the role of the Gibbs Ensemble and entropy maximization in statistical mechanics; however, it is ultimately a logical statement about statistical inference and, as such, statistical thermodynamics is just one application of the Principle of Maximum Entropy.

The articulation of the Principle of Maximum Entropy was made possible by the development of a theory of information and a (recovered) understanding that probability is a more general concept than “relative frequency of occurrence”. That a probability distribution embodies information is a relatively straightforward proposition to demonstrate. Consider a physical system that can be in any one of several states and suppose that we can measure certain properties μ, ν, \dots of the system whose values depend on — but do not necessarily determine — the system’s state. For example, the system might be an N -dimensional quantum simple harmonic oscillator, whose state is determined by the N quantum numbers n_1, n_2, \dots, n_N , and the observable might be the total energy of the system, $\mu = \hbar(N/2 + \sum_k n_k)$. If we have no knowledge of the system’s state then it is natural that we should regard it as equally likely that the system is in any particular state: i.e., if we were to assign a number p , $0 \leq p \leq 1$, to represent our degree of certainty that the system’s state is \vec{n} , with 0 representing complete certainty that the system is not in state \vec{n} and 1 representing complete certainty it is in state \vec{n} , then p would be the same for all \vec{n} . Now suppose that we measure or are told μ is equal to μ_0 : what does this new information tell us about the state of the system? In our example, we now know that only those states \vec{n} satisfying $\sum_k n_k$ equal to $\mu_0/\hbar - N/2$ are possible but that, among these, no \vec{n} is preferred over another: i.e., the number p we should ascribe to state \vec{n} vanishes if $\mu(\vec{n})$ is not equal to μ_0 and is equal to a constant for all other states. Given this distribution we can express our expectations regarding other properties of the system: e.g., our expectation that the property ν takes on the value ν_0 is equal to the

³In fact, the noise cross-spectral density — equivalent to the correlation of the noise between two different detectors — has not yet been evaluated for or played a role in the analysis of LIGO data.

ratio of the number of states for which (ν, μ) is equal to (ν_0, μ_0) to the number of states for which μ is equal to μ_0 . These numbers p both represent our degree of certainty regarding the state of the system *and* satisfy exactly the same algebraic and relational laws that we require of probabilities (Cox 1946, 1961); so, we will hereafter refer to them as probabilities.

In this demonstration we see that probability is more general than “relative frequency of occurrence”, that it does not require that we invoke an ensemble or make any assumptions regarding ergodicity — there is just one system and it is in a definite state — and that it represents in a meaningful way our state of knowledge about the system. Modern information theory was launched in a series of papers by Shannon (1948a,b) in which he gave a definite meaning to “information content” in the context of a probability distribution, described the properties that a measure of information should have, and showed that there exists a unique measure of the information content of a probability distribution, which is the negative of what is conventionally understood as the entropy of an ensemble. It is more common to refer to the negative of the information content as the measure of uncertainty, or (information) entropy:

$$H(p) = - \sum_k p_k \log p_k \quad (4)$$

where the sum is over all states of the system.⁴ For a continuous distribution $p(\vec{x})$ the summation becomes an integral and we must normalize the probability as it occurs in the logarithm by the density of states $\rho(\vec{x})$:

$$H(p) = - \int d^n x p(\vec{x}) \log \frac{p(\vec{x})}{\rho(\vec{x})}. \quad (5)$$

Returning to our previous example, it is clear that there may be a large number of distributions $q(\vec{n})$ that are consistent with the knowledge that the observed μ is equal to μ_0 . Each of these distributions has an entropy $H(q)$; of these, one — call it \tilde{q} — will have the greatest entropy. We can find that distribution by finding the q such that the variation $\delta H(q)/\delta q$, subject to the constraint that $\sum_k q_k = 1$, vanishes:

$$\begin{aligned} 0 &= \delta \left[H(q) + \lambda_0 \sum_k q_k \right] \\ &= \sum_k [-1 - \log q_k + \lambda_0] \delta q_k \end{aligned} \quad (6a)$$

$$q_k = \exp(\lambda_0 - 1) \quad (6b)$$

⁴The base of the logarithm corresponds simply to a choice of units in which to measure entropy.

where the Lagrange multiplier λ_0 is chosen so that $\sum_k q_k = 1$. The maximum entropy distribution \tilde{q} assigns equal probability to every state \vec{n} such that $\mu(\vec{n})$ is equal to μ_0 : i.e., \tilde{q} is the distribution that embodies what we understand to be just the information that $\mu(\vec{n})$ is equal to μ_0 , but no more. The Principle of Maximum Entropy states that this is always so: i.e., given a system with possible states x and constraints on the system in the form of functions of the state, the probability distribution $p(x)$ that embodies just the information in those constraints is equal to the distribution that maximizes the entropy (either equation 4 or 5) subject to the constraints.

We can use the Maximum Entropy Principle to find the likelihood function g — i.e., the probability distribution that $\mathbf{d} - \mathbf{R}\mathbf{h}$ is equal to detector noise — given the characterization of the gravitational wave detector noise described by its mean μ and autocorrelation function c_ℓ :

$$\mu = \langle x \rangle \sim \lim_{N \rightarrow \infty} N^{-1} \sum_{k=0}^{N-1} x_k \quad (7a)$$

$$c_\ell = \langle x_k x_{k+\ell} \rangle \sim \lim_{N \rightarrow \infty} N^{-1} \sum_{k=0}^{N-1} (x_k - \mu)(x_{k+\ell} - \mu). \quad (7b)$$

We can construct many example probability densities $p(\mathbf{x})$ whose moments take on these values. From among these we desire the particular probability distribution $\tilde{p}(\vec{x})$ that has maximum entropy: i.e., we desire the distribution that has minimum information subject to the constraints that its moments satisfy equations 7. This distribution satisfies the variational equation

$$0 = \delta \left[- \sum_{\vec{x}} \tilde{p}(\vec{x}) \log \tilde{p}(\vec{x}) - \lambda_0 \left(1 - \sum_{\vec{x}} \tilde{p}(\vec{x}) \right) - \lambda_1 \left(\mu - \sum_{\vec{x}} x_k \tilde{p}(\vec{x}) \right) - \sum_{\ell} \lambda_{2,\ell} \left(c_\ell - \sum_{\vec{x}} (x_k - \mu)(x_{k+\ell} - \mu) \tilde{p}(\vec{x}) \right) \right] \quad (8)$$

where λ_0 , λ_1 and $\lambda_{2,\ell}$ are Lagrange coefficients, which are chosen to insure that the probability is normalized (λ_0), that the mean $\langle x_k \rangle$ is μ (λ_1), and that the auto-correlation $\langle x_k x_{k+\ell} \rangle$ takes on the value c_ℓ ($\lambda_{2,\ell}$). Solving equation 8 is straightforward: taking the variation we find

$$0 = \sum_{\vec{x}} \left[\left(1 + \log \tilde{f}(\vec{x}) \right) + \lambda_0 + \lambda_1 x_k + \sum_{\ell} \lambda_{2,\ell} \sum_{\vec{x}} (x_k - \mu)(x_{k+\ell} - \mu) \right]; \quad (9)$$

i.e., $\log \tilde{f}$ is quadratic in \vec{x} , subject to the constraints of unitarity and equations 7, or

$$\tilde{f}(\vec{x}) = \left[(2\pi)^{\dim \vec{x}} \det ||C|| \right]^{-1/2} \exp \left[-\frac{1}{2} (\vec{x} - \vec{\mu})^T C^{-1} (\vec{x} - \vec{\mu}) \right] \quad (10)$$

where $\vec{\mu}$ is the dim \vec{x} vector all of whose elements are μ and the covariance matrix C^{-1} is the related to the autocorrelation function by

$$C_{jk} = c_{k-j}. \quad (11)$$

When nothing more than the noise distributions mean and covariance are known the least presumptive model for the noise statistics is a multivariate Gaussian. Adopting any other noise model makes additional (and, since we know only μ and c_ℓ , unjustified) assumptions. This important result is contrary to popular prejudice, which views Gaussian noise models suspiciously even when nothing more is available for use in analysis than the noise mean and covariance. Correspondingly, for the analysis of gravitational wave detector data we have

$$\log g(\mathbf{d}|\mathbf{h}', \mathbf{R}, \mathcal{N}, \mathcal{I}) = -\frac{1}{2}\chi^2(\mathbf{d}, \mathbf{h}', \mathbf{R}, \mathbf{N}) + \text{const.} \quad (12)$$

where

$$\chi^2(\mathbf{d}, \mathbf{h}', \mathbf{R}, \mathbf{N}) = (\mathbf{d} - \mathbf{R}\mathbf{h}')^T \mathbf{N}^{-1} (\mathbf{d} - \mathbf{R}\mathbf{h}') \quad (13)$$

may be evaluated directly: i.e., without the need to invert the generally ill-conditioned response \mathbf{R} .

2.3. The a prior probability q

When the observables \mathbf{d} over-determine the incident wave \mathbf{h} it is tempting to ignore $q(\mathbf{h}'|\mathcal{I})$ in equation 3, minimize χ^2 (i.e., maximize the likelihood) over \mathbf{h}' , and declare that \mathbf{h}' is the inferred incident wave. This is, with minor variation, the approach taken by Gürsel & Tinto (1989). While an entirely legitimate approach to deconvolution, maximum likelihood methods generally over-fit \mathbf{d} : i.e., they find a \mathbf{h}' that leaves a residual $\mathbf{d} - \mathbf{R}\mathbf{h}'$ that is inconsistent with the known noise properties \mathcal{N} . One role played by $q(\mathbf{h}'|\mathcal{I})$ is to resist this tendency toward over-fitting noisy data.

One commonly thinks of the a priori probability $q(\mathbf{h}'|\mathcal{I})$ as a representation of our general expectation signals \mathbf{h} prior to the particular observation \mathbf{d} . Criticisms of Bayesian analyses generally focus on the necessity of invoking a priori probabilities like q . A great deal of intellectual energy has been devoted to developing ways of identifying a priori probabilities that are, in some meaningful sense, without prejudice. The Maximum Entropy Principle, as described above, provides one means of doing so, which is especially useful when the set of system states is discrete or the density of states is known. Another approach, also pioneered by Jaynes (1968), takes as its starting point the idea that the prior q should be equivalent for equivalent experiments: more formally, that q should be form-invariant under the same set of

transformations that leave the experiment invariant. Such priors encode only our knowledge of the experiment’s nature and do not presuppose any outcome. We can use that principle in the present instance to find the class of priors q that are the least presumptive regarding the signal \mathbf{h} .

Let us first focus on some general considerations that should govern the prior q . Writing \mathbf{h} as the sum of two time series, corresponding to the two polarizations \mathbf{h}_+ and \mathbf{h}_\times of an incident gravitational wave, we note that

- Lacking any reason to presume that the source is oriented in a particular way relative to the detector line-of-sight, the prior should be invariant under an arbitrary rotation of $+$ into \times polarization;
- Lacking any reason to presume that the gravitational wave burst arrives at a particular moment in time, the prior for $h(t)$ and $h(t + \tau)$ should be the same⁵; and, finally,
- Lacking any specific source model that dictates or suggests how energy is distributed throughout the burst, the prior should not favor waveforms with any particular degree of smoothness (or autocorrelation or spectrum).

We could, of course, introduce specific knowledge or assumptions into the construction of the prior, if we have any: for example, we may believe that the gravitational wave power should be concentrated in a band of frequencies, or that there is a relationship between the radiation in the $+$ and \times polarization states, etc. In the present instance, however, our goal is to specify a prior that makes the fewest possible assumptions (beyond propagation direction, which we take to be known) about the nature of the gravitational wave signal incident on our detector(s).

With these general considerations in mind we may write \mathbf{h}_+ (\mathbf{h}_\times) in a Fourier expansion:

$$h_k = \sum_j A_j \cos(\omega_j t_k) + B_j \sin(\omega_j t_k). \quad (14)$$

The prior on \mathbf{h} can just as well be expressed as a prior Q on the \vec{A} and \vec{B} . From our general considerations and the property of Fourier series we conclude that

$$Q(\vec{A}, \vec{B} | \mathcal{I}) = \prod_k \bar{Q}(A_k | \mathcal{I}) \bar{Q}(B_k | \mathcal{I}); \quad (15)$$

⁵In the case of galactic core-collapse supernova we expect that the arrival time of the observed neutrinos will coincide with the arrival time of the gravitational wave burst up to uncertainties in the time they spend trapped in the dense core. Within these uncertainties we treat the arrival time of the gravitational wave burst as unknown

i.e., the overall prior is the product of priors of identical form evaluated for each of the coefficients A_k and B_k .

Bretthorst (1988) provides a particularly lucid derivation, which we reproduce here, of the prior \bar{Q} that arises from the transformation group properties that we expect it should satisfy. To begin, we might just as well have written the Fourier expansion of \mathbf{h} in terms of the Fourier amplitudes a_j and phases θ_j , as in

$$h_k = \sum_j a_j \cos(\omega_j t_k + \theta_j), \quad (16)$$

with the (a_j, θ_j) related to the (A_j, B_j) by

$$a_j^2 = A_j^2 + B_j^2 \quad (17a)$$

$$\tan \theta_j = B_j / A_j. \quad (17b)$$

The same considerations that led us to factor $Q(\vec{A}, \vec{B}|\mathcal{I})$ into the product of identical functions \bar{Q} of the A_k and B_k lead us to factor the prior $P(\vec{a}, \vec{\theta}|\mathcal{I})$ into the product of priors \bar{P} of identical form, evaluated for each of the coefficient pairs (a_k, θ_k) :

$$P(\vec{a}, \vec{\theta}|\mathcal{I}) = \prod_k \bar{P}(a_k, \theta_k|\mathcal{I}) \quad (18)$$

Under the assumption that we do not know the signal arrival time all phases θ are equally likely: i.e.,

$$\bar{P}(a, \theta|\mathcal{I}) = \frac{\bar{p}(a|\mathcal{I})}{2\pi}. \quad (19)$$

The choice of Fourier coordinates (A_k, B_k) or (a_k, θ_k) does not affect our prior knowledge; so, the prior distributions \bar{P} and \bar{Q} must be equivalent: i.e.,

$$\bar{Q}(A|\mathcal{I})\bar{Q}(B|\mathcal{I})dA dB = \frac{\bar{p}(a|\mathcal{I})}{2\pi}da d\theta. \quad (20)$$

In the special case where B is equal to zero this relation becomes

$$\bar{p}(a|\mathcal{I})da d\theta = 2\pi\bar{Q}(a|\mathcal{I})\bar{Q}(0|\mathcal{I})dA dB. \quad (21)$$

Using this expression for \bar{p} in equation 20 we find that the desired prior \bar{Q} must satisfy the functional equation

$$\bar{Q}(x|\mathcal{I})\bar{Q}(y|\mathcal{I}) = \bar{Q}\left(\sqrt{x^2 + y^2}\right)\bar{Q}(0|\mathcal{I}), \quad (22)$$

which has the general solution

$$\bar{Q}(x|\sigma^2, \mathcal{I}) = \frac{\exp[-x^2/2\sigma^2]}{\sqrt{2\pi\sigma^2}} \quad (23)$$

for some unknown parameter σ^2 .

We are thus led to a one parameter family of “uninformative” priors $q_\sigma(\mathbf{h}|\mathcal{I})$:

$$q_\sigma(\mathbf{h}|\mathcal{I}) = \prod_k \tilde{q}_\sigma(A_k|\mathcal{I})\tilde{q}_\sigma(B_k|\mathcal{I}) \quad (24a)$$

$$= \frac{\exp\left[-\frac{1}{2\sigma^2}\sum_{k=0}^{N-1}(A_k^2 + B_k^2)\right]}{(2\pi\sigma^2)^{N/2}} \quad (24b)$$

$$= \frac{\exp\left[-\frac{1}{2\sigma^2}\sum_{k=0}^{N-1}h_k^2\right]}{(2\pi\sigma^2)^{N/2}}. \quad (24c)$$

Setting aside for the moment the question of how we might choose σ , the choice of prior q taking the form $q_\sigma(\mathbf{h}|\mathcal{I})$ leads us to the posterior probability density

$$f(\mathbf{h}'|\mathbf{d}, \mathbf{R}, \mathcal{N}, \sigma, \mathcal{I}) \propto g(\mathbf{d}|\mathbf{h}', \mathbf{R}, \mathcal{N}, \mathcal{I})q(\mathbf{h}'|\sigma, \mathcal{I}) \quad (25)$$

$$\propto \exp\left[-\frac{1}{2}\chi^2(\mathbf{h}', \mathbf{R}, \mathcal{N}, \mathbf{d}) + \frac{1}{2\sigma^2}S(\mathbf{h}')\right] \quad (26)$$

where

$$S(\mathbf{h}') = -\sum_k h_k'^2. \quad (27)$$

We may thus choose as the best estimate of \mathbf{h} the \mathbf{h}' that minimizes $\ln f$, or

$$F(\mathbf{h}'|\mathbf{d}, \mathbf{R}, \mathcal{N}, \sigma^2) = \frac{1}{2}\chi^2(\mathbf{h}', \mathbf{R}, \mathcal{N}, \mathbf{d}) - \frac{1}{2\sigma^2}S(\mathbf{h}). \quad (28)$$

Stripped of its Bayesian statistical motivation, we recognize that maximum entropy deconvolution of the gravitational signal \mathbf{h} is closely related to the solution of the inverse problem via regularization, with the specific choice $\sum_k h_k^2$ for the regularization function S (Neumaier 1998). Nevertheless, our Bayesian approach to inference is much more than a motivation for a particular choice of regularization function, or even for deconvolution via regularization. Whereas deconvolution provides a single estimate of the signal \mathbf{h} , our Bayesian approach provides a distribution f , whose mode is the point estimate of the equivalent deconvolution problem but that also assigns a probability to any proposed \mathbf{h}' . The availability of this distribution permits, among other things, a detailed error analysis for any problem where the details of the waveform play an important role.

2.4. Eliminating the regularization constant

If we view the determination of \mathbf{h} from \mathbf{d} as a problem in deconvolution via regularization, then we require some prescription for the choice of regularization constant σ^2 . Viewed

as a problem of Bayesian inference, however, σ^2 is not unlike any other unknown and we'd prefer, instead, to find some probability distribution $u(\sigma^2)$ that describes the probability that σ^2 is the appropriate choice and integrate the q_σ over u to find a distribution q independent of σ^2 : i.e.,

$$q(\mathbf{h}|\mathcal{I}) = \int d\sigma^2 u(\sigma^2)q(\mathbf{h}|\sigma^2, \mathcal{I}) \quad (29)$$

Suppose we let the data guide us to a choice of $u(\sigma^2)$: i.e., we use Bayes' Theorem to construct the probability distribution u of σ^2 , based on the observations \mathbf{d} :

$$u(\sigma^2|\mathbf{d}, \mathcal{N}, \mathbf{R}, \mathcal{I}) \propto \int D\mathbf{h} g(\mathbf{d}|\mathbf{h}', \mathbf{R}, \mathcal{N}, \mathcal{I})q(\mathbf{h}'|\sigma^2, \mathcal{I})t(\sigma^2|\mathcal{I}) \quad (30)$$

where we have introduced the prior $t(\sigma^2|\mathcal{I})$.

In general we have no knowledge that suggests a preferred value of σ^2 ; correspondingly, any reasonable prior t should be approximately constant over a wide range of σ^2 . Additionally, when the data \mathbf{d} are informative then gq_σ will be a sharply peaked (relative to t) function of σ^2 . For both these reasons we may treat t as constant over the support of gq_σ as a function of σ^2 in the integrand of equation 29, in which case

$$u(\sigma^2|\mathbf{d}, \mathcal{N}, \mathbf{R}, \mathcal{I}) \propto \int D\mathbf{h} g(\mathbf{d}|\mathbf{h}', \mathbf{R}, \mathcal{N}, \mathcal{I})q_\sigma(\mathbf{h}'|\mathcal{I}) \quad (31)$$

This integral we recognize as the normalization constant v that appears in our expression for f (cf. eq. 3) when evaluated for a particular value of σ . *In addition to allowing us to infer the signal \mathbf{h} our observations \mathbf{d} also provide us with a probability distribution for the “regularization constant” σ^2 .*

For our problem the distribution $u(\sigma^2|\mathbf{d}, \mathcal{N}, \mathbf{R}, \mathcal{I})$ involves only Gaussian integrals and is relatively straightforward to calculate: recalling our definition of χ^2 (cf. eq. 13) and introducing

$$Z_g = \int D\mathbf{d} \exp\left(-\frac{1}{2}\chi^2\right) = \left[(2\pi)^{N_d} \det ||\mathbf{N}||\right]^{1/2} \quad (32a)$$

$$Z_q = \int D\mathbf{h} \exp\left(-\frac{\mathbf{h}^2}{2\sigma^2}\right) = (2\pi\sigma^2)^{N_h/2} \quad (32b)$$

$$Z_f = \int D\mathbf{h} \exp\left(-\frac{1}{2}\chi^2 - \frac{\mathbf{h}^2}{2\sigma^2}\right) = \left[\frac{(2\pi)^{N_h}}{\det ||\sigma^{-2}\mathbf{I} + \mathbf{R}^T\mathbf{N}^{-1}\mathbf{R}||}\right]^{1/2} \quad (32c)$$

we find

$$\begin{aligned} 2 \log u &= -\chi^2(\mathbf{d}, \mathbf{R}\mathbf{h}_0, \mathcal{N}) - \sigma^{-2}S(\mathbf{h}_0) + 2 \log \frac{Z_f}{Z_g Z_q} \\ &= -\chi_0^2 - \frac{S_0}{\sigma^2} - N_d \log 2\pi + 2N_h \log \sigma - \log \det ||\sigma^{-2}\mathbf{I} + \mathbf{R}^T\mathbf{N}^{-1}\mathbf{R}||, \end{aligned} \quad (32d)$$

where \mathbf{h}_0 is the \mathbf{h} that maximizes $g(\mathbf{h})q_\sigma(\mathbf{h})$, N_d is $\dim \mathbf{d}$, and N_h is $\dim \mathbf{h}$

We can generally approximate equation 29 for $q(\mathbf{h}|\mathcal{I})$ by $q_{\hat{\sigma}}(\mathbf{h}|\mathcal{I})$, where $\hat{\sigma}^2$ maximizes $u(\sigma^2|\mathcal{I})$. While conceptually we do not make a choice of σ , as a matter of practice if a signal is recoverable f depends at most weakly on q and the desired prior q is also generally close to $q_{\hat{\sigma}}$. To find $\hat{\sigma}$ we can extremize our expression for $\log u$ (cf. 29) over σ :

$$0 = \frac{d \log u}{d\sigma} = \frac{S_0}{\sigma^3} - \frac{N_h}{\sigma} + \sigma^{-3} \text{tr} \left[\left(\sigma^{-2} \mathbf{I} + \mathbf{R}^T \mathbf{N}^{-1} \mathbf{R} \right)^{-1} \right] \quad (33a)$$

or

$$\hat{\sigma}^2 = \frac{\sum_k h_k^2}{N_h - \text{tr} \left[\left(\mathbf{I} + \hat{\sigma}^2 \mathbf{R}^T \mathbf{N}^{-1} \mathbf{R} \right)^{-1} \right]}, \quad (33b)$$

where we made use of the relation

$$\ln \det \|\mathbf{A}\| = \text{tr} \ln \|\mathbf{A}\|. \quad (33c)$$

Equation 33b provides an implicit relationship for $\hat{\sigma}^2$, which can be solved iteratively with the minimization of equation 28 for \mathbf{h} : e.g., we can solve equation 28 for \mathbf{h} given a guess for $\hat{\sigma}^2$ and use this $(\hat{\sigma}^2, \mathbf{h})$ pair on the right-hand side of equation 33b to find a new estimate for $\hat{\sigma}^2$, with the process repeated until it converges.

2.5. Example

As an example, consider a plane gravitational wave signal propagating from a hypothetical source directly overhead of the LIGO Hanford Observatory (LHO) site, and incident on detectors at the LIGO Hanford and LIGO Livingston observatory sites. For this example assume that

- Each site has a single, identical interferometric gravitational wave detector oriented in the same way as the actual detectors at each site;
- The response function for each detector is frequency independent in the band of interest;
- The detector noise is independent between detectors and white (i.e., the autocorrelation is a delta function in lag), with unit variance in each detector;
- The sample rate at each detector is 4096 Hz.

For the signal use

$$\mathbf{h} = h_+ \mathbf{e}_+ + h_\times \mathbf{e}_\times \quad (34a)$$

$$h_+ = h_0 \cos \phi_0 e^{-(t-t_0)^2/2\sigma_t^2} \cos[2\pi f(t-t_0)] \quad (34b)$$

$$h_\times = h_0 \sin \phi_0 e^{-(t-t_0)^2/2\sigma_t^2} \sin[2\pi f(t-t_0)] \quad (34c)$$

$$|t - t_0| < 24.4 \text{ ms} \quad (34d)$$

$$f = 414 \text{ Hz} \quad (34e)$$

$$\sigma_t = 4 \text{ ms} \quad (34f)$$

with gravitational wave polarization tensor \mathbf{e}_\times orthogonal to tangents to lines of latitude, \mathbf{e}_+ orthogonal to \mathbf{e}_\times , and ϕ_0 is chosen so signal projection on the LHO detector is maximized.

Figure 1 shows snippets of the simulated data from five different data sets \mathbf{d} . Each data set corresponds to the (noisy) observation at the detector network described above when a signal of amplitude h_0 equal to 1, 2, 5, 10, or 20, is incident on the detector network. Each snippet is centered on the location of the actual signal, which was 100 samples in duration and embedded in a data set five times as long. Figure 2 shows the most probable waveforms h'_+ and h'_\times as inferred, in the manner just described, from each data set. Finally, figure 3 shows the residual $\mathbf{d} - \mathbf{R}\mathbf{h}'$ for each data set and inferred waveform \mathbf{h}' . Table 1 shows the power signal-to-noise ratio ρ^2 of the recovered signals, calculated as⁶

$$\rho^2 = (\mathbf{R}\mathbf{h}')^T \mathbf{N}^{-1} (\mathbf{R}\mathbf{h}') = \chi^2(0, \mathbf{h}', \mathbf{R}, \mathbf{N}) \quad (35)$$

and the correlation between the actual and the inferred signal, calculated as

$$C(j) = \frac{\sum_k h_{k+j} h'_k}{|\mathbf{h}'| |\mathbf{h}|} \quad (36a)$$

where

$$|\mathbf{x}|^2 = \sum_k x_k^2. \quad (36b)$$

The value of ρ^2 and $C(j)$ recorded here should be taken as representative: their actual values depend on the noise instantiation, which varies from simulation to simulation (and from observation to observation).

⁶It is important to note that the signal-to-noise as defined here involves only the inferred signal: i.e., it is the observed signal-to-noise, which varies from simulation to simulation with the different instantiations of the noise. Calculated in this way it is a different quantity than the signal-to-noise often quoted in the context of experimental results published by the LIGO Scientific Collaboration, who report their sensitivity in terms of the *expectation value* of the signal-to-noise for a *known* signal — not an inferred one — supposed to be present in the data (see, for example Abbott et al. (2004b)).

Comparing figures 1 and 2 we see the degree to which our inference procedure is able to separate signal from noise. For the cases h_0 equal to 1 or 2 no signal is resolved: i.e., h_0 is uniformly small and, in fact, much smaller than the rms noise. For the three cases h_0 equal to 5, 10 and 20 the inference procedure clearly does resolve distinct signals in each of the two gravitational wave polarization. The maximum cross-correlations (cf. equation 36), tabulated in table 1, makes the correctness of the inference as a function of h_0 quantitative: the correlation between the real and inferred signals is greater than a half for signal amplitudes greater than 5, and inspection of figure 2 shows that the error is principally in the signal amplitude and not its time dependence.

Should we expect to be able to do better? Estimates of the required power signal-to-noise for reliable detection of a gravitational wave burst in LIGO based solely on the coincidence of excess power in the detector data streams⁷ range from 30 to 70 (Thorne 1987; Finn 1991, 1992). Referring to table 1 it is apparent that a reliable signal emerges very rapidly between signal amplitudes h_0 of 2 and 5, corresponding to ρ^2 in this range. Figure 3 shows $\mathbf{R}\mathbf{h}' - \mathbf{d}$, the residual after the inferred signal is subtracted from the observations. If we have done a good job of separating the signal and noise contributions to the observations \mathbf{d} than we expect that $\mathbf{R}\mathbf{h}' - \mathbf{d}$ should be noise: in this example, unity rms white noise. In the cases h_0 equal to 1 or 2 inspection of the residuals shows no obvious evidence of structure left behind by the inference procedure. For the case h_0 equal to five and greater, where a clear signal is identified, the residual does show some evidence of structure near the location of the signal peak. The amplitude of this apparent structure is very much consistent with our assumptions about the detector noise (i.e., unity rms): it is only the time dependence of the structure that draws the attention of the eye. Over longer random time series, even this would not appear to be particularly unusual: i.e., were we not aware that we were looking at a residual we would not have reason based on statistics to be particularly suspicious of the presence of a signal here. Reviewing the construction of our inference procedure, the quantitative agreement between our estimates of the required signal-to-noise for a reliable detection, and the qualitative evidence of these residuals, it is clear that absent additional information about the nature of the noise (which would modify our function χ^2) or the nature of the signal (which would modify our choice of function S) we cannot reasonably suppose that there is any further information in the residuals about a possible signal.

⁷And calculated in a manner comparable to the ρ^2 recorded in table 1

2.6. Summary

The problem of identifying a signal in noisy data can be approached as a problem of statistical inference: i.e., we can find a probability distribution that describes the credibility that we should ascribe to the hypothesis that an arbitrary gravitational wave burst \mathbf{h} was incident on an array of detectors given observations \mathbf{d} . This probability distribution involves two components: the probability g that $\mathbf{d} - \mathbf{Rh}$ is detector noise, and the a priori probability q that a signal takes the form \mathbf{h} .

The probability g , also known as the likelihood function, requires a characterization of the noise. In the absence of a complete characterization of the noise — something that is rarely, if ever, possible — the Maximum Entropy Principle provides a means of selecting the probability g that is simultaneously consistent with what is known about the noise and least presumptive about what is not. In particular, when the noise is characterized by its mean and covariance — the usual case — then g takes the form of a Normal distribution.

The a priori probability density q characterizes our prior expectations regarding signals \mathbf{h} . In the spirit of “letting the data speak for itself” we avoid prejudicing our analysis with preconceived notions regarding the nature of potential sources by focusing instead on our ignorance of signal polarization, arrival time, and energy spectrum. In this way we are led to a one-parameter family of potential priors q_σ . Rather than choose a particular prior from this family, we find that the data allow us to provide probabilistic weights $u(\sigma)$ to different values of the parameter σ . Rather than make a particular choice of distribution q_σ , we evaluate the expectation value of q_σ over the distribution of the parameter σ and use it for q , eliminating our dependence on the unknown parameter σ . As a practical matter, it is often sufficient to find the $\hat{\sigma}$ that maximizes $u(\sigma)$ and use $q_{\hat{\sigma}}$ for q .

3. APPLICATION TO LIGO OBSERVATIONS OF SN

How well can we hope to characterize the astrophysics of core-collapse supernovae from the gravitational wave signature we observe? The detailed astrophysics of core-collapse supernovae is uncertain, difficult to model, and involves large-scale convection and other stochastic processes. For all these reasons, it is more likely that we will infer the features of the gravitational wave burst associated with a supernova and then use that inferred signal to validate our models, than that we will “discover” the gravitational wave signal from a core-collapse supernova by using matched filtering to extract it from deep within a noisy data stream.

As a first step toward exploring how, and how well, we may be able to diagnose the con-

ditions of the collapsing core from the gravitational waves it radiates we use the techniques, described in the previous two sections, to infer a simulated signal embedded in LIGO-like noise and evaluate the cross-correlation between the inferred signal and a wide range of signals calculated from the models of Ott et al. (2004). We do this for

- LIGO data drawn from the S1–S4 science runs, evaluating the maximum distance at which there is a significant correlation between the inferred and simulated signals (cf. sec. 3.2.1);
- Simulated signals with varying maximum central density (cf. sec. 3.2.2), evaluating our ability to evaluate the properties of the core-collapse;
- Simulated signals with varying progenitor mass, evaluating our ability to test the correlation between progenitor models and the configuration of the collapsing stellar core (cf. sec. 3.2.3);
- Simulated signals with varying angular momentum and differential angular velocity, allowing us to explore the role that differential rotation plays in core collapse (cf. sec. 3.2.4).

We find that, even when the simulated signal is so weak that the inferred signal appears very different, the inferred signal still has its greatest cross-correlation with the simulated signal. We conclude that the model whose signal has the maximum cross-correlation with the inferred signal is likely to provide a good indication of the physical properties of the source.

3.1. The Gravitational Wave Signature of Core-Collapse Supernovae: Description of Core-Collapse Models

Gravitational wave emission from core-collapse supernovae may arise from a multitude of processes, including rotating core collapse and core bounce (e.g., Fryer et al. 2002; Dimmelmeier et al. 2002; Ott et al. 2004; Dimmelmeier et al. 2007a,b; Ott et al. 2007), post-bounce convective overturn, anisotropic neutrino emission (Burrows & Hayes 1996; Müller et al. 2004; Ott et al. 2007), nonaxisymmetric rotational instabilities of the protoneutron star (Rampp et al. 1998; Ott et al. 2007; Shibata & Sekiguchi 2005), or from the recently proposed protoneutron star core g-mode oscillations (Burrows et al. 2006, 2007; Ott et al. 2006; Ferrari et al. 2003). In addition and in the context of the core-collapse supernova – gamma-ray burst connection (Woosley & Bloom 2006), late-time black hole formation in a failed or

weak core-collapse supernova may accompany gravitational wave emission from quasi-normal ring-down modes of the newly-formed black hole.

Of all the above emission processes, rotating iron core collapse and bounce is the most extensively modeled and best quantitatively and qualitatively understood. For these reasons we limit our present study to an analysis of the rotating core collapse and bounce signature only and use example templates computed from a set of 2D axisymmetric Newtonian core collapse simulations by Ott et al. (2004) which focused on the dynamics of rotational collapse and bounce. These represent a class of signatures of core collapse and are good templates with which to exercise the signal extraction technology we have developed. This paper is the first to put theoretical models of the gravitational wave signals of core collapse and bounce through a realistic detector pipeline and to attempt to extract physical information using sophisticated signal processing algorithms.

The astrophysics models involved stellar progenitors with various masses: 11, 15, 20, and $25M_{\odot}$ calculated in Woosley & Weaver (1995). The simulations neglected the effects of neutrinos, general relativity, and magnetic fields, but used the realistic, finite-temperature nuclear equation of state of Lattimer & Swesty (1991). A small number of simulations also investigated stellar progenitor models from Heger et al. (2000) and Heger et al. (2004), which were evolved to the onset of iron core collapse with an approximate treatment of rotation (Heger et al. 2000; Heger et al. 2004) and angular momentum redistribution by magnetic torques (Heger et al. 2004). The gravitational wave signature extraction was performed using the Newtonian quadrupole formalism (see e.g. Misner et al. 1973).

The effects of rotation were investigated in Ott et al. (2004). The initial rotation of the progenitor was controlled by two parameters: the rotation parameter β where

$$\beta = \frac{E_{\text{rot}}}{|E_{\text{grav}}|}, \quad (37)$$

and the differential rotation scale parameter A , which is the distance from the rotational axis at which the rotational velocity drops to half that at the center. A is defined as

$$\Omega(r) = \Omega_0 \left[1 + \left(\frac{r}{A} \right)^2 \right]^{-1}, \quad (38)$$

where r is the distance from the axis of rotation and $\Omega(r)$ is the angular frequency at r .

When the progenitor is rotating slowly and β is small (zero to a few tenths of a percent), the collapse is halted when the inner core reaches supranuclear densities. The core bounces rapidly and then quickly rings down. When the progenitor rotates more rapidly and β is larger, the core collapse is halted by centrifugal forces and the core bounces at subnuclear

densities. The core then undergoes multiple damped, harmonic oscillator-like expansion-collapse-bounce cycles⁸ The initial degree of differential rotation affects the value of β at which this bounce type transition occurs. A progenitor with a smaller value of A experiences a greater amount of differential rotation and hence, a more rapidly rotating inner core. As a result, the transition from a supranuclear to a subnuclear bounce occurs for a lower value of β .

The models of Ott et al. (2004) yield absolute values of the dimensionless maximum gravitational wave strain h_{\max} in the interval $2 \times 10^{-23} \leq h_{\max} \leq 1.25 \times 10^{-20}$ at a detector distance of 10 kpc. The total energy radiated (E_{GW}) lies in the range $1.4 \times 10^{-11} M_{\odot}c^2 \leq E_{\text{GW}} \leq 2.21 \times 10^{-8} M_{\odot}c^2$ and most of it is emitted in the primary gravitational wave burst associated with core bounce. The energy spectra peak in the frequency interval $20 \text{ Hz} \leq f_{\text{peak}} \leq 600 \text{ Hz}$ with rapid and differential rotators having peaks at low frequencies and moderate and rigid rotators peaking at high frequencies.

3.2. Gravitational wave observations of rotating stellar core collapse

In all the simulations described here we used Ott et al. (2004) model s15A1000B0.1 as the “real” signal, and correlated the inferred signal with other models in the Ott et al. (2004) catalog. Model s15A1000B0.1 corresponds to a $15 M_{\odot}$ Woosley & Weaver (1995) progenitor with rotation parameter β equal to 0.1% and differential rotation scale parameter A equal to 1000 km. We scaled this signal to represent core collapse events at different distances and projected the incident signal onto the LIGO 4-km Hanford WA (LHO) and Livingston, LA (LLO) detectors. For the purpose of this study we assumed that the core collapse was directly overhead of the LHO site.⁹ Since the Ott et al. (2004) core collapse models are axisymmetric the gravitational waves they emit are linearly polarized. We chose the polarization angle to maximize the response of the LHO detector and used the actual response functions for the LHO and LLO detectors characteristic of LIGO’s S4 science run (González et al. 2004). We simulated the detector noise by adding white noise with power spectral density amplitude approximately equal to the noise amplitude at 100 Hz in the

⁸Recent results of Ott et al. (2007) and Dimmelmeier et al. (2007a,b) suggest that such multiple-bounce dynamics are less likely when general relativity and deleptonization are taken into account.

⁹In this work we have always assumed that the gravitational wave signal we observe is accompanied by a neutrino or electromagnetic signal that determines the source sky position.

corresponding science run (Lazzarini 2005).¹⁰ Finally, we used the maximum entropy method described above to find the best point estimate of the embedded signal and cross-correlated this estimate signal with different signals drawn from the Ott et al. (2004); Ott et al. (2007) parameter survey. The subsections below describe our observations based on this study.

3.2.1. *Science Run and Survey Range*

Before discussing how well we can distinguish between different core collapse models we looked at how close a core collapse event would have needed to be for LIGO, during its first four science runs (S1, S2, S3 and S4), to infer its waveform with reasonable accuracy. For this purpose we used the detector response functions and noise power spectral density amplitudes for the corresponding science run (Adhikari et al. 2003; González et al. 2004; Lazzarini 2002, 2003, 2004, 2005).

Figure 4 shows the maximum cross-correlation, for each of the first four science runs, between the inferred and actual waveforms as a function of the core collapse distance. There is a steady improvement, from S1 to S4, in maximum entropy’s ability to recover signals at greater distances, corresponding to improving detector sensitivity: by S4, we are able to infer the gravitational waveform from core collapse events that occur as far as a few kpc away. The up-tick in cross-correlation at the 1 kpc mark in curve for S3 sensitivity is the result of a discontinuous change in the most probable waveform (i.e., a new extrema becoming the global maximum in the probability function.) The investigations described in the following sections, which delve into the source information present in the inferred signal, use the response and noise characteristic of the LIGO S4 science run.

An alternative way of expressing the results summarized in Figure 4 is to describe the sensitivity required of a LIGO-like detector to observe an optimally located supernovae to a given distance. Under the “white noise” approximation made here, doubling the distance to which LIGO is able to infer the waveform from a core collapse supernova requires doubling its sensitivity: i.e., halving the rms noise amplitude at 100 Hz. Referenced to LIGO’s design sensitivity at 100 Hz, which was achieved during LIGO’s S5 science run, and a 70% maximum

¹⁰At the frequencies where the signal power (in units of squared strain) of the simulated core collapse event gravitational wave signals peaks (~ 500 Hz) the noise power spectral density in any of the current generation of interferometric gravitational wave detectors is increasing relatively slowly with frequency. (Very little is known concerning the cross-correlation in the noise of two separated detectors, other than that it is very small.) While we have chosen to make this demonstration with white noise, the maximum entropy analysis method introduced above naturally accommodates any noise covariance: cf. the paragraph including equation (12).

cross-correlation, the maximum distance at which core collapse supernova are expected to be observable is

$$d_{\max}(70\%) = 4 \text{ kpc} \left(\frac{3 \times 10^{-23} \text{ Hz}^{-1/2}}{\sqrt{S_h(100 \text{ Hz})}} \right) \quad (39)$$

where $S_h(f)$ is the strain-equivalent noise power spectral density. The intermediate LIGO upgrade underway at this writing is expected to reduce S_h by a factor of four, and the advanced LIGO upgrade that will follow will reduce S_h by another factor of approximately 100 (LIGO 2007). Correspondingly, we expect that advanced LIGO will be capable of observing core-collapse supernovae like those modeled here at distances as great as 80 kpc: i.e., beyond the distance to the LMC and SMC, but far short of the Virgo cluster.

3.2.2. Bounce Type

We can classify the models of Ott et al. (2004) into those that bounce at supranuclear, subnuclear and transitional central densities. Figure 5 shows the maximum cross-correlation between the inferred waveform, the actual waveform (a supranuclear bounce type), and three other waveforms that each have, within their respective categories (supernuclear, subnuclear and transitional), the greatest cross-correlation with the inferred waveform. (Figure 6 shows these four waveforms themselves, each with strain scaled to a distance of 10 kpc.) It is clear that the inferred waveform has the most in common with that generated from a model with the same, supranuclear, bounce type and that, for S4 detector sensitivities, our ability to distinguish bounce type fails for core collapse events more than 3 kpc distant.

To facilitate comparison with signal-to-noise ratio sensitivities as reported by the LIGO Scientific Collaboration for the LIGO detectors, we show the quantity “SNR²”, defined by

$$\text{SNR}^2 = \frac{1}{2} (H_1^2 + H_2^2), \quad (40a)$$

$$H_k^2 = \sum_{j=0}^N \frac{(\mathbf{r}_k \mathbf{h})^2}{N \sigma_k^2}, \quad (40b)$$

where h is the *actual* waveform, running across the top of this and subsequent figures. The quantity SNR² can be compared to the single-detector, optimal orientation, expectation value of the power signal-to-noise ratio sensitivities reported by LIGO in its publications: i.e., assuming that the actual gravitational waveform were known, the expectation value of the signal-to-noise determined by matched filtering would be given by SNR.

3.2.3. Mass

Figure 7 shows the cross-correlation between the inferred and actual waveforms associated with core collapse models that differ by progenitor mass, but share the same rotational parameters. Figure 8 shows the waveforms used for the cross-correlations. Again, for S4 sensitivities the inferred waveform most closely resembles the waveform from the model with the same mass at distances up to 2-3 kpc.

3.2.4. Rotation

Figure 9 shows the cross-correlation between the inferred waveform, the actual waveform, and the waveforms associated with a set of models that differ only by rotation parameter β , while figure 10 shows the cross-correlations for models that differ only by the differential rotation parameter A (cf. eq. 38). Figures 11 and 12 show the waveforms for the set of models whose cross-correlations are shown in figures 9 and 10. As expected, the cross-correlation decreases as the rotational parameters of the models depart from those associated with the actual waveform.

4. DISCUSSION/CONCLUSIONS

We have described and demonstrated the use of a Bayesian method, in the spirit of the so-called Maximum Entropy methods for analyzing astronomical image data, for inferring the time-dependent waveforms of the radiation incident on a network of gravitational wave detectors. In contrast to the methods developed for image analysis, which rely on the Principle of Maximum Entropy to select an a priori probability for use in the Bayesian analysis, we use the MaxEnt Principle to determine the appropriate choice of likelihood function for describing gravitational wave detector noise and use transformation group arguments to identify an appropriate, uninformative a priori probability density. We have demonstrated that the method described is quite capable of inferring the presence of signal in detector noise once the detector response to the signal begins to approach the noise rms, and very quickly converges on the correct signal amplitude and temporal structure of the incident signal as the actual signal amplitude increases. We have argued, based on the methods construction, a quantitative assessment of the correlation between the inferred and actual signal and the observed and expected signal-to-amplitudes associated of simulated data sets, and on a qualitative review of the residuals following subtraction of the detector networks response to the *inferred* signal from the observations themselves, that the method developed here is the best

that one can expect to be able to do for detector networks whose noise character is known only via its power spectral density and for signals that are assumed to have arbitrary time dependence.

Inferring the gravitational wave signal incident on a set of detectors is but the first task of analysis. Our ultimate goal is to use the inferred signal to understand the radiation source. We demonstrate how this might be done using gravitational wave signatures calculated from the rotating core collapse supernovae models of Ott et al. (2004). Using simulations drawn from Ott et al. (2004) we have generated simulated LIGO observations of core collapse supernovae, used our Bayesian analysis method to infer the most probable gravitational wave signal content of the data, and correlated this inferred signal with a wide range of signals drawn from the catalog of Ott et al. (2004). We have found, in the context of this example, that at the signal-to-noise ratios expected to be used as a threshold for detection of core-collapse supernova burst events in LIGO, the correlation is maximized when the inferred signal is matched against its parent from the Ott et al. (2004) catalog: thus, the analysis method we describe here infers the incident gravitational wave signal with enough fidelity to distinguish between models of different maximum central density, angular momentum, differential rotation, and progenitor mass.

To make the comparison between the inferred and model waveforms we used the maximum of the cross-correlation between the two, choosing the model whose maximum cross-correlation with the inferred waveform is the greatest. A Bayesian treatment that assigns odds to different models given the inferred waveform is a natural next step; however, we note that a Bayesian treatment would involve the exponential of the cross-correlation function (as opposed to the maximum cross-correlation) marginalized over the lag. When the cross-correlation is a sharply peaked function, this will be proportional to the exponential of the maximum of the cross-correlation among the different models. Thus, we conclude that the cross-correlation as we use it here is almost certainly likely to choose the same model as a more sophisticated Bayesian analysis, except in those cases where the waveforms of the different models are nearly indistinguishable.

While our example applications have been in terms of two interferometers, there is nothing in the description of the method that specifies the number of detectors and the application to three or more detectors is no different than for one or two detectors. The example application provided here involved just the two main LIGO detectors; however, we could just as well have augmented the analysis by including simulated data Virgo, GEO, TAMA and the operating bar detectors. For the particular simulations described here, adding any one of these detectors would most likely not improve our estimates of the signal since — for the examples presented here — we chose to place the source in the most favorable sky

location and orientation for detection by the LIGO Hanford Observatory (LHO) site. For a source so positioned the strain amplitude in the LIGO Livingston detector is 89% of that for the LHO detectors, while for the GEO600 site detector it is 42%, 22% for TAMA300, and 1% for Virgo. On the other hand, the addition of these other detectors to a real network would increase the overall sensitivity across the sky, as a source optimally oriented for Virgo would, by reciprocity, be very poorly oriented for the LHO detectors.

Gravitational wave astronomy is not yet a reality. Nevertheless, with the detectors now operating and the enhancements currently underway, it is only a matter of time before we have observations from these instruments, from which we can gain insight into astronomical phenomena that are otherwise hidden from our sight. Methods like those described here will allow us to move beyond gravitational wave detection and realize the full promise of gravitational wave astronomy.

We are grateful to our anonymous referee for helpful recommendations, which have helped us significantly strengthen this paper's accessibility and relevance. Tzs and Lsf gratefully acknowledge informative discussions with Graham Woan, Malik Rakhmanov, Nelson Christensen, and Soumya Mohanty. Tzs and Lsf were supported by NSF through grant PHY-00-99559 and the Center for Gravitational Wave Physics. Lsf also acknowledges the support of NSF through awards PHY 06-53462 and PHY 05-55615, NASA through award NNG05GF71G, and the Penn State Center for Space Research Programs. AB acknowledges support for this work from the Scientific Discovery through Advanced Computing (SciDAC) program of the DOE, grant numbers DE-FC02-01ER41184 and DE-FC02-06ER41452 and from the NSF under grant number AST-0504947. CDO acknowledges support through a Joint Institute for Nuclear Astrophysics postdoctoral fellowship, sub-award no. 61-5292UA of NFS award no. 86-6004791. The core-collapse calculations were carried out on the local Arizona cluster, at the Albert Einstein Institute, and at the National Center for Supercomputing Applications (NCSA) under Teragrid computer time grant TG-MCA02N014. The Center for Gravitational Wave Physics is funded by the National Science Foundation via cooperative agreement PHY 01-14375.

REFERENCES

Abbott, B., Abbott, R., Adhikari, R., Ageev, A., Allen, B., Amin, R., Anderson, S. B., Anderson, W. G. et al. 2004a, Nucl. Instrum. Methods Phys. Research, 517, 154

Abbott, B., Abbott, R., Adhikari, R., Ageev, A., Allen, B., Amin, R., Anderson, S. B., Anderson, W. G. et al. 2004b, Phys. Rev. D, 69, 122001

Acernese, F., Amico, P., Alshourbagy, M., Antonucci, F., Aoudia, S., Avino, S., Babusci, D., Ballardin, G. et al. 2006, *Class. Quantum Grav.*, 23, S635

Adhikari, R., González, G., Landry, M., & O'Reilly, B. 2003, *Class. Quantum Grav.*, 20, S903

Barreiro, R. B., Hobson, M. P., Banday, A. J., Lasenby, A. N., Stolyarov, V., Vielva, P., & Górski, K. M. 2004, *MNRAS*, 351, 515

Barreiro, R. B., Vielva, P., Hobson, M. P., Martínez-González, E., Lasenby, A. N., Sanz, J. L., & Toffolatti, L. 2001, in *Mining the Sky*, ed. A. J. Bandy, S. Zaroubi & M. Bartelmann (Berlin: Springer-Verlag), 465

Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L. et al. 2003, *ApJS*, 148, 1

Bretthorst, G. L. 1988, in *Lecture Notes in Statistics*, Vol. 48, *Bayesian Spectrum Analysis and Parameter Estimation*, ed. J. Berger, S. Fienberg, J. Gani, K. Krickeberg, & B. Singer (New York: Springer-Verlag)

Burrows, A., & Hayes, J. 1996, *Phys. Rev. Lett.*, 76, 352

Burrows, A., Livne, E., Dessart, L., Ott, C. D., & Murphy, J. 2006, *ApJ*, 640, 878

—. 2007, *ApJ*, 655, 416

Cox, R. T. 1946, *Am. Jour. Phys.*, 14, 1

—. 1961, *The Algebra of Probable Inference* (Baltimore: Johns Hopkins Press)

Dimmelmeier, H., Font, J. A., & Müller, E. 2002, *A&A*, 393, 523

Dimmelmeier, H., Ott, C. D., Janka, H.-T., Marek, A., & Mueller, E. 2007a, ArXiv e-prints, 705

Dimmelmeier, H., Ott, C. D., Janka, H.-T., Marek, A., & Müller, E. 2007b, *Phys. Rev. Lett.*, 98, 251101

Evans, S. N., & Stark, P. B. 2002, *Inverse Problems*, 18, R55

Ferrari, V., Miniutti, G., & Pons, J. A. 2003, *Class. Quantum Grav.*, 20, 841

Finn, L. S. 1991, *Annals Of The New York Academy Of Sciences*, 631, 156

—. 1992, *Phys. Rev. D*, 46, 5236

Finn, L. S., & Chernoff, D. F. 1993, Phys. Rev. D, 47, 2198

Fryer, C. L., Holz, D. E., & Hughes, S. A. 2002, ApJ, 565, 430

González, G., Landry, M., O'Reilly, B., & Radkins, H. 2004, Calibration of the LIGO detectors for S2, Tech. Rep. T040060-01-D, Laser Interferometer Gravitational Wave Observatory, LIGO Document Control Center, California Institute of Technology, available at <http://admdbsrv.ligo.caltech.edu/dcc>

Gull, S. F., & Daniell, G. J. 1978, Nature, 272, 686

Gull, S. F. & Daniell, G. J. 1979, in Astrophysics and Space Science Library, Vol. 76, Image Formation from Coherence Functions in Astronomy, ed. C. van Schooneveld, IAU (C. D. Reidel), 219

Gürsel, Y., & Tinto, M. 1989, Phys. Rev. D, 40, 3884

Heger, A., Langer, N., & Woosley, S. E. 2000, ApJ, 528, 368

Heger, A., Woosley, S. E., Langer, N., & Spruit, H. C. 2004, in IAU Symposium, Vol. 215, Stellar Rotation, ed. A. Maeder & P. Eenens (San Francisco: Astronomical Society of the Pacific), 591–+

Jaynes, E. T. 1957a, Phys. Rev., 106, 620

—. 1957b, Phys. Rev., 108, 171

—. 1968, IEEE Transactions on Systems Science and Cybernetics, 4, 227

Jones, A. W., Hancock, S., Lasenby, A. N., Davies, R. D., Gutierrez, C. M., Rocha, G., Watson, R. A., & Rebolo, R. 1998, MNRAS, 294, 582

Jones, A. W., Lasenby, A. N., Mukherjee, P., Gutierrez, C. M., Davies, R. D., Watson, R. A., Hoyland, R., & Rebolo, R. 1999, MNRAS, 310, 105

Kittel, C. 1958, Elementary Statistical Physics (New York: Wiley)

Lattimer, J. M., & Swesty, F. D. 1991, Nucl. Phys A, 535, 331

Lazzarini, A. 2002, LIGO S1 Strain Sensitivity, Tech. Rep. G020461-01-E, Laser Interferometer Gravitational Wave Observatory, LIGO Document Control Center, California Institute of Technology, available at <http://admdbsrv.ligo.caltech.edu/dcc>

- . 2003, LIGO S2 Strain Sensitivity, Tech. Rep. G030379–00-E, Laser Interferometer Gravitational Wave Observatory, LIGO Document Control Center, California Institute of Technology, available at <http://admdbsrv.ligo.caltech.edu/dcc>
- . 2004, S3 Best Strain Sensitivities, Tech. Rep. G040023–00-E, Laser Interferometer Gravitational Wave Observatory, LIGO Document Control Center, California Institute of Technology, available at <http://admdbsrv.ligo.caltech.edu/dcc>
- . 2005, S4 Strain Sensitivities for the LIGO Interferometers, Tech. Rep. G050230–02-E, Laser Interferometer Gravitational Wave Observatory, LIGO Document Control Center, California Institute of Technology, available at <http://admdbsrv.ligo.caltech.edu/dcc>

LIGO Laboratory. 2007, <http://ligo.caltech.edu>

Lück, H., Hewitson, M., Ajith, P., Allen, B., Aufmuth, P., Aulbert, C., Babak, S., Balasubramanian, R. et al. 2006, *Class. Quantum Grav.*, 23, 71

Maisinger, K., Hobson, M. P., & Lasenby, A. N. 1997, *MNRAS*, 290, 313

—. 2004, *MNRAS*, 347, 339

Mandic, V. 2006, PoS, HEP2005, 037

Misner, C. W., Thorne, K. S., & Wheeler, J. A. 1973, *Gravitation* (San Francisco: W. H. Freeman & Co.)

Müller, E., Rampp, M., Buras, R., Janka, H.-T., & Shoemaker, D. H. 2004, *ApJ*, 603, 221

Narayan, R., & Nityananda, R. 1984, in *Indirect Imaging. Measurement and Processing for Indirect Imaging*, ed. J. A. Roberts (Cambridge, England: Cambridge University Press), 281

Narayan, R., & Nityananda, R. 1986, *ARA&A*, 24, 127

Neumaier, A. 1998, *SIAM Review*, 40, 636

Nityananda, R., & Narayan, R. 1982, *Journal of Astrophysics and Astronomy*, 3, 419

—. 1983, *A&A*, 118, 194

Ott, C. D., Burrows, A., Dessart, L., & Livne, E. 2006, *Phys. Rev. Lett.*, 96, 201102

Ott, C. D., Burrows, A., Livne, E., & Walder, R. 2004, *ApJ*, 600, 834

Ott, C. D., Dimmelmeier, H., Marek, A., Janka, H.-T., Hawke, I., Zink, B., & Schnetter, E. 2007, Phys. Rev. Lett., 98, 261101

Ott, C. D., Dimmelmeier, H., Marek, A., Janka, H.-T., Zink, B., Hawke, I., & Schnetter, E. 2007, Class. Quantum Grav., 24, 139

Pantin, E., & Starck, J.-L. 1996, Astron. Astrophys. Supp., 118, 575

Ponsonby, J. E. B. 1973, MNRAS, 163, 369

Rakhmanov, M. 2006, Class. Quantum Grav., S673

Rampp, M., Mueller, E., & Ruffert, M. 1998, A&A, 332, 969

Robertson, N. A. 2000, Class. Quantum Grav., 17, 19

Rüdiger, A. 2004, in Observing, Thinking and Mining the Universe, ed. G. Miele & G. Longo (Singapore: World Scientific Press), 329–+

Shannon, C. E. 1948a, The Bell System Technical Journal, 27, 327

—. 1948b, The Bell System Technical Journal, 27, 623

Shevgaonkar, R. K. 1987, A&A, 176, 159

Shibata, M., & Sekiguchi, Y.-I. 2005, Phys. Rev. D, 71, 024014

Skilling, J., & Bryan, R. K. 1984, MNRAS, 211, 111

Starck, J., & Pantin, E. 1996, Vistas in Astronomy, 40, 563

Starck, J. & Pantin, E. 1997, in Statistical Challenges in Modern Astronomy II, ed. G. J. Babu & E. D. Feigelson (Berlin: Springer-Verlag), 405

Steenstrup, S. 1985, Australian Journal of Physics, 38, 319

Takahashi, R., & the TAMA Collaboration. 2004, Class. Quantum Grav., 21, S403

Thorne, K. S. 1987, in Three Hundred Years of Gravitation, ed. S. Hawking & W. Israel (Cambridge, U.K.; New York, U.S.A.: Cambridge University Press), 330–458

Vielva, P., Barreiro, R. B., Hobson, M. P., Martínez-González, E., Lasenby, A. N., Sanz, J. L., & Toffolatti, L. 2001, MNRAS, 328, 1

Waldman, S. J., & The LIGO Scientific Collaboration . 2006, Class. Quantum Grav., 23, S653

Woosley, S. E., & Bloom, J. S. 2006, *ARA&A*, 44, 508

Woosley, S. E., & Weaver, T. A. 1995, *ApJS*, 101

Table 1: Signal to noise ρ^2 and maximum cross-correlation $\max_j C(j)$ for a simulated gravitational wave signal incident on a two-detector LIGO-like network. The signal and detector network model are described in §2.5 .

h_0	ρ^2	$\max_j C(j)$
1	0.29	0.08
2	1.5	0.11
5	1.5×10^2	0.54
10	1.0×10^3	0.83
20	5.4×10^3	0.94

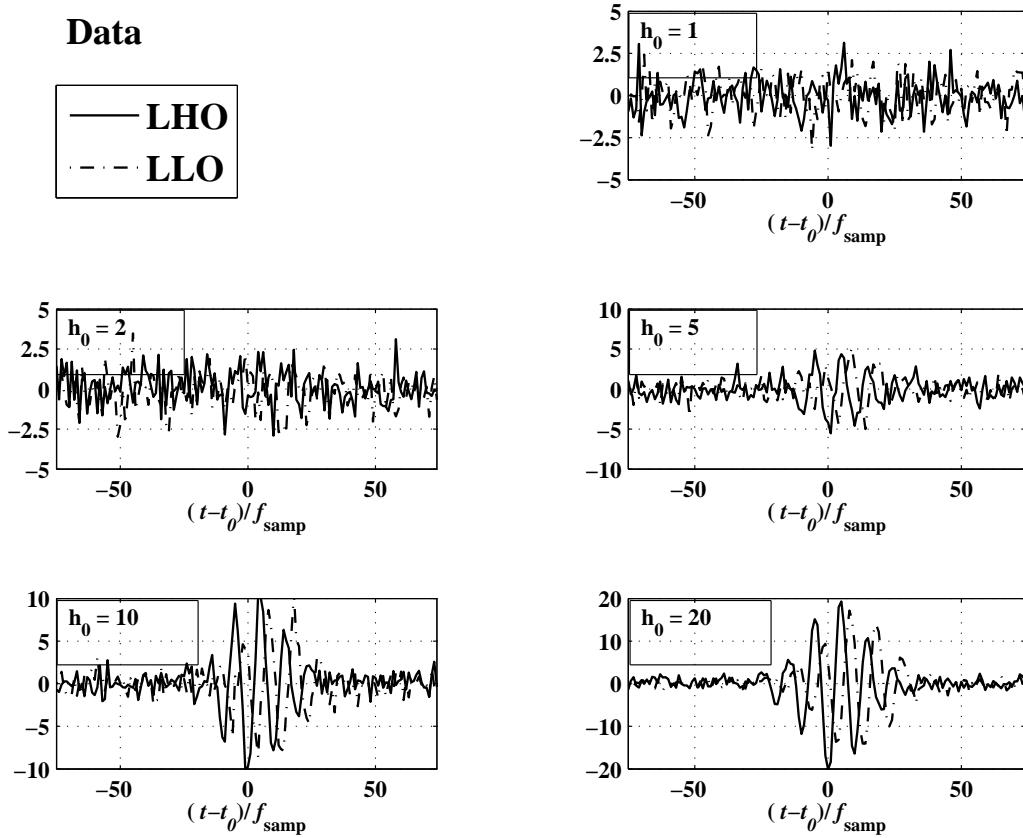


Fig. 1.— A short snippet of the simulated data surrounding the embedded signal in the simulations described in §2.5. The embedded signal was approximately 100 samples long and centered in the snippet shown. The demonstration analysis was carried out on a data segment five times longer than the signal. See §2.5 for more details.

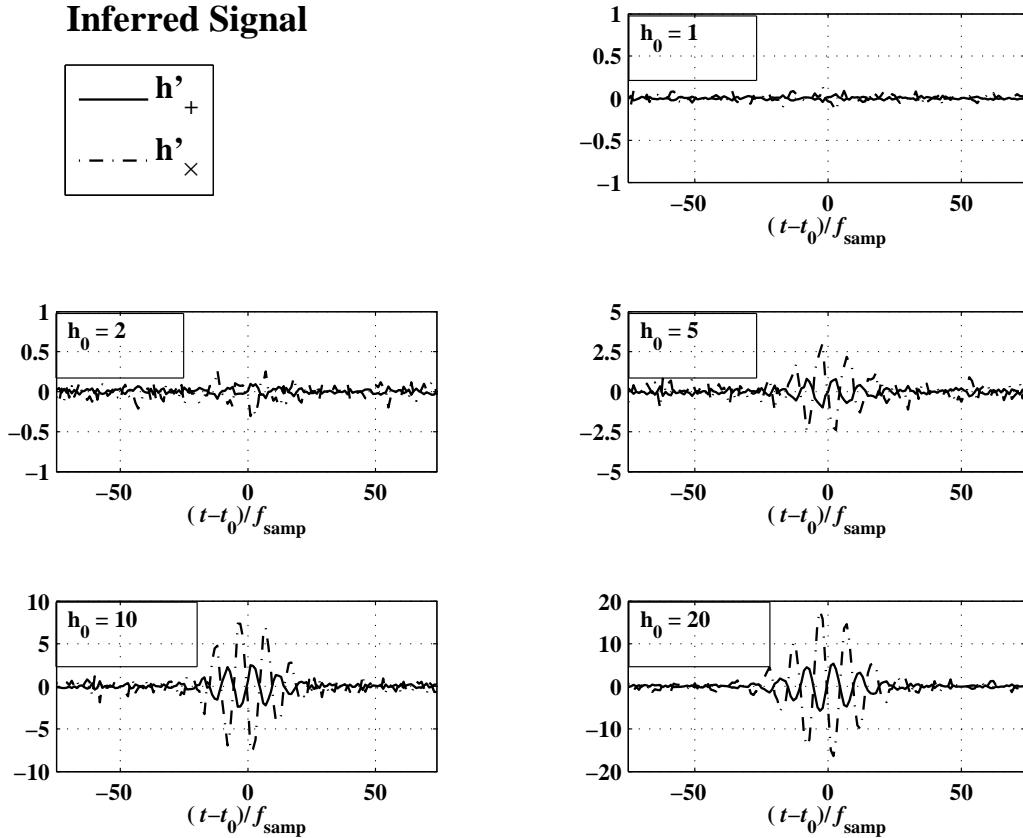


Fig. 2.— The inferred gravitational wave signal in the + and \times polarization states for the five data sets described in §2.5. Each panel here corresponds to the data shown in the corresponding panel of figure 1. Note that the two detector network does an excellent job of rejecting detector noise (which has unity variance) and that the actual signal emerges very rapidly once the signal amplitude begins to exceed the noise amplitude. See §2.5 for more details.

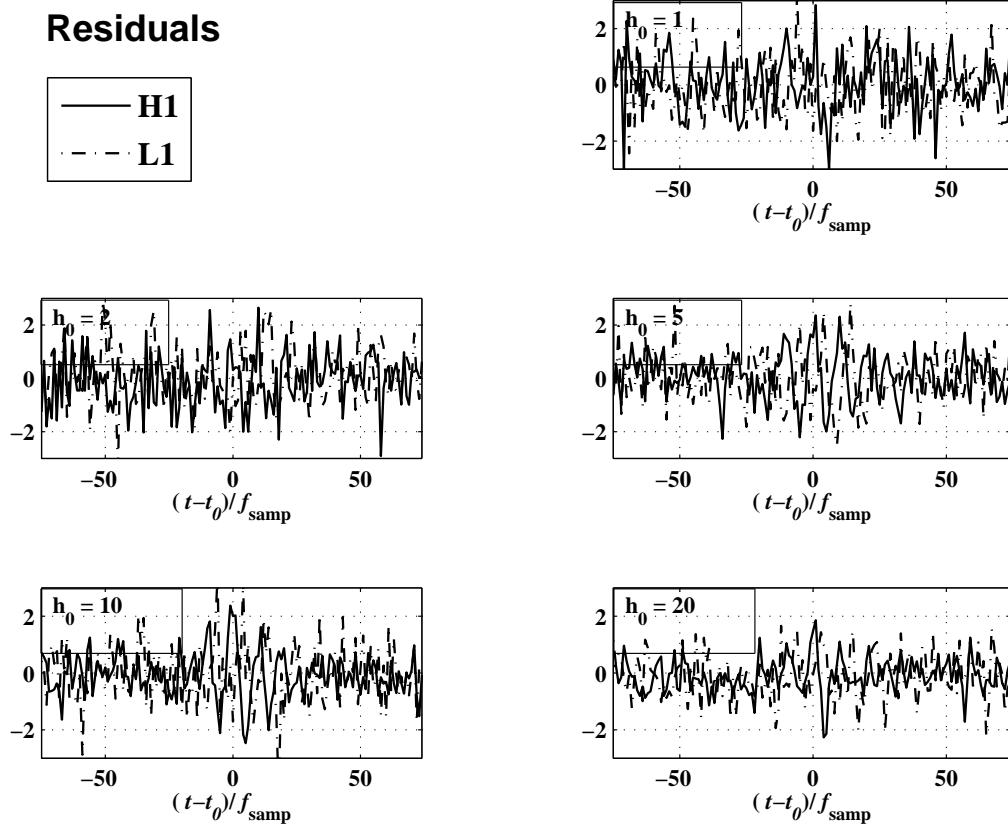


Fig. 3.— The residuals $\mathbf{d} - \mathbf{R}\mathbf{h}'$ following subtraction of the response of the simulated detector network to the inferred signal ($\mathbf{R}\mathbf{h}'$) from the observations (\mathbf{d}) made in the network, as described in §2.5. Each panel here corresponds to the data shown in the corresponding panel of figures 1 and 2. The amplitude of the residuals is consistent with detector noise and the temporal structure of the residuals is, in the absence of additional knowledge regarding either the actual signal or the detector noise, also consistent with the detector noise. See §2.5 for more details.

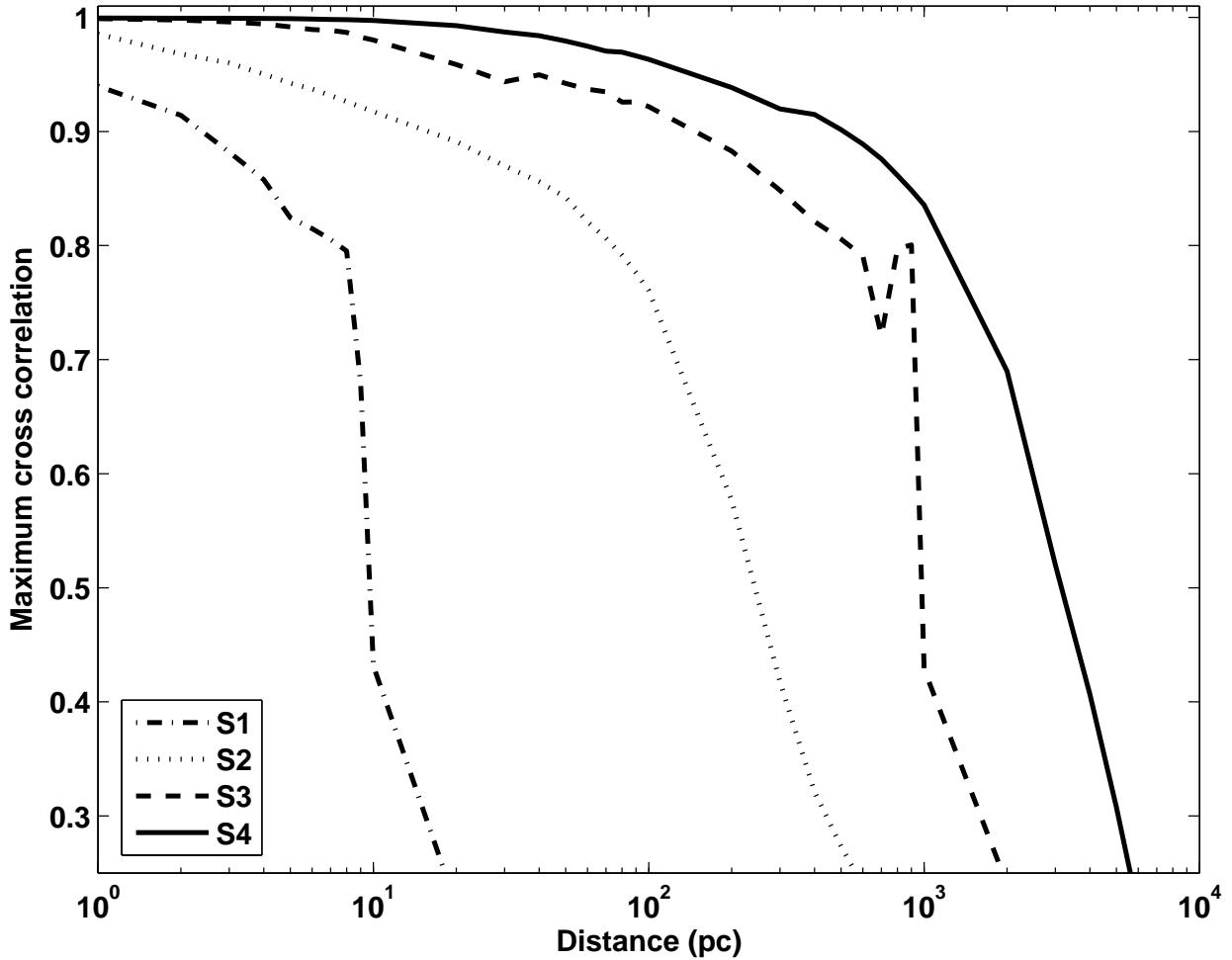


Fig. 4.— Maximum cross-correlation between estimated and initial signals versus core collapse distance for data simulated using different science run detector impulse responses and noise levels. The estimated signal is inferred using maximum entropy from simulated detections that use Ott et al. (2004) model s15A1000B0.1 as the initial signal waveform. The up-tick in the S3 run, located at approximately 1 kpc, is the result of a discontinuous change in the most probable waveform. There is a steady improvement in maximum entropy’s ability to reconstruct fainter, more distant signals as the sensitivity of the detectors improved (Lazzarini 2002, 2003, 2004, 2005).

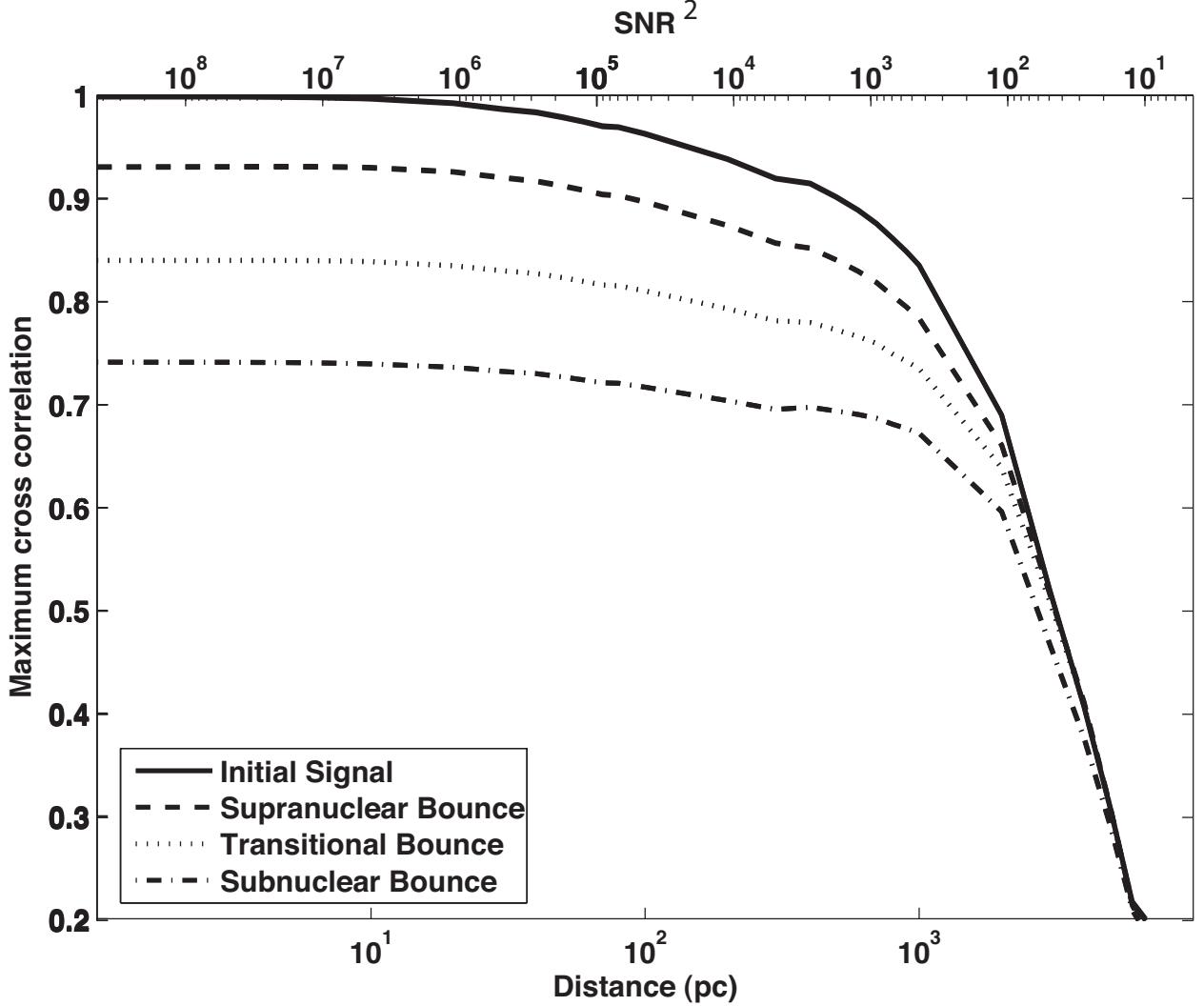


Fig. 5.— Maximum cross-correlation between reconstructed waveforms and waveforms associated with models that differ by bounce type versus core collapse distance and SNR^2 . The SNR^2 shown is the average for the two detectors. The SNR^2 for the 4-km Hanford detector is 1.05 times that shown while the SNR^2 for the 4-km Livingston detector is 0.95 times that shown. The reconstructed waveform is inferred using maximum entropy from simulated detections that use the waveform from the Ott et al. (2004) model s15A1000B0.1 as the initial signal waveform as well as detector responses and noise levels from the fourth science run (S4). The solid line represents the maximum cross-correlation between the reconstructed signal and the initial signal waveform. The other lines represent the maximum cross-correlations between the inferred waveforms and the waveforms resulting from each bounce type for which the maximum cross-correlation at 1 pc is greatest, excluding that used for the initial signal. The inferred waveform is most similar to those generated by models with the same, supranuclear bounce type as the initial signal waveform, for the simulations corresponding to core collapse events that occur less than 2-3 kpc away.

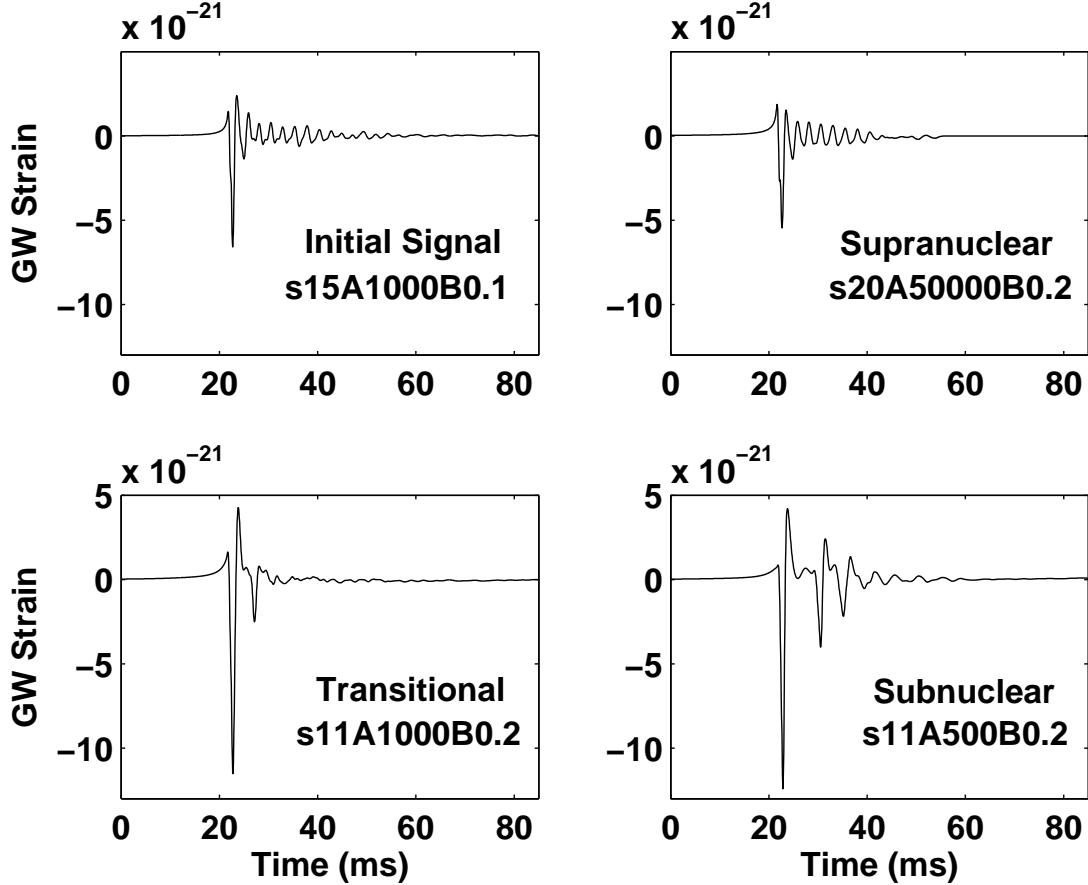


Fig. 6.— The waveforms associated with various bounce types that are compared with the inferred signal in Fig. 5. The upper left plot shows the waveform (from Ott et al. [2004] model s15A1000B0.1) that was used as the initial signal in the simulated detection. The three other waveforms shown are those that are most similar to this initial signal waveform, for each bounce type. The waveform from the Ott et al. (2004) s20A50000B0.2 model looks much like the initial signal waveform which is of the same, supranuclear bounce type. The subnuclear bounce waveform shows the effects of multiple damped, harmonic oscillator-like expansion-collapse-bounce cycles. The zero points of the time axes are chosen so that the minima of the waveforms occur at the same time for ease of comparison. The waveform amplitudes are scaled to correspond to core collapse events at 10 kpc.

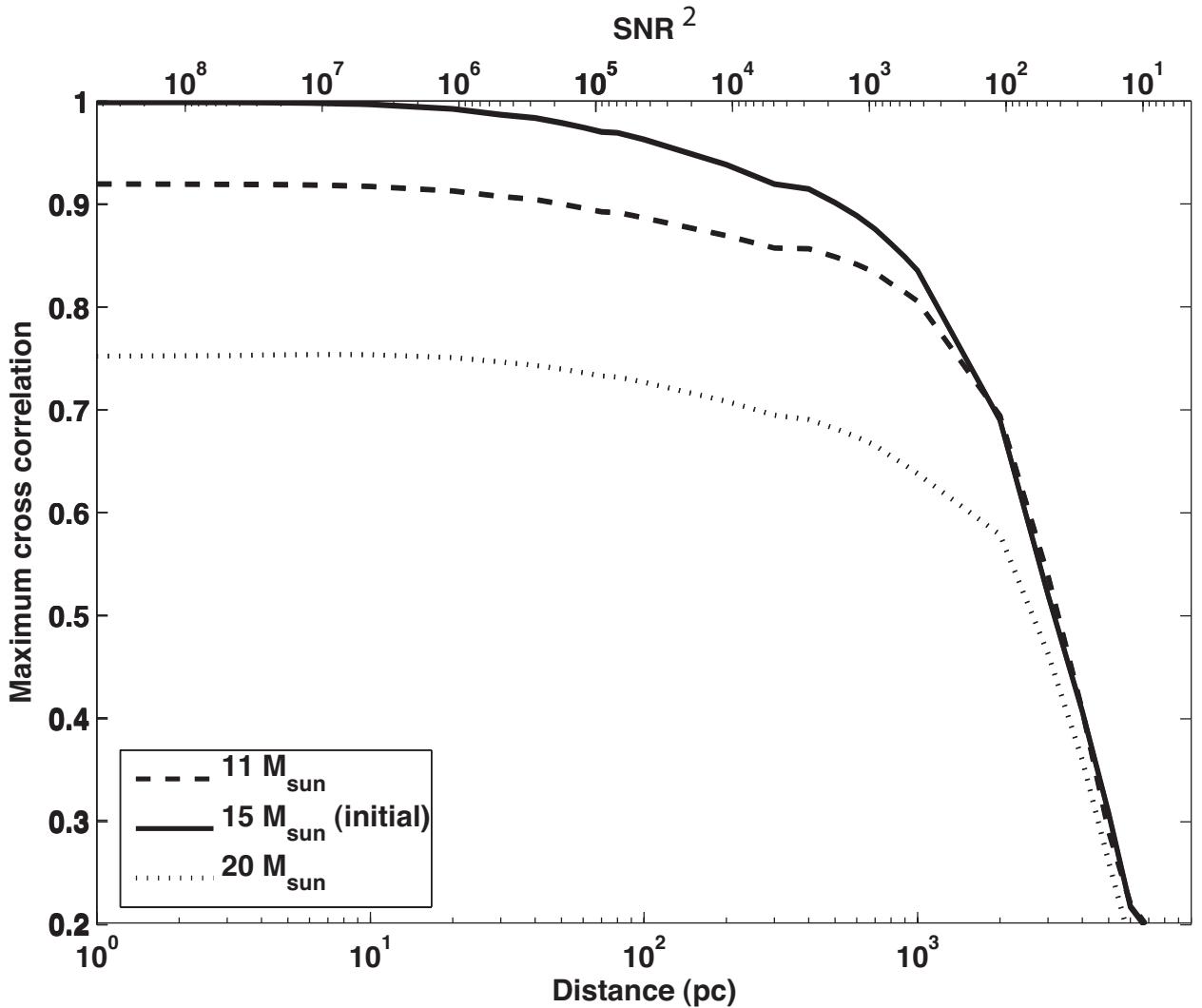


Fig. 7.— Maximum cross-correlation between reconstructed waveforms and waveforms associated with models that differ only by progenitor mass versus core collapse distance and SNR^2 . The SNR^2 shown is the average for the two detectors. The SNR^2 for the 4-km Hanford detector is 1.05 times that shown while the SNR^2 for the 4-km Livingston detector is 0.95 times that shown. The reconstructed waveform is inferred using maximum entropy from simulated detections that use a waveform from a model with a progenitor mass of 15 solar masses (Ott et al. [2004] model s15A1000B0.1) as the initial signal waveform as well as detector responses and noise levels from the fourth science run (S4). The inferred waveform is most similar to that generated by the model with the same progenitor mass for the simulations corresponding to core collapse events that occur less than 2-3 kpc away.

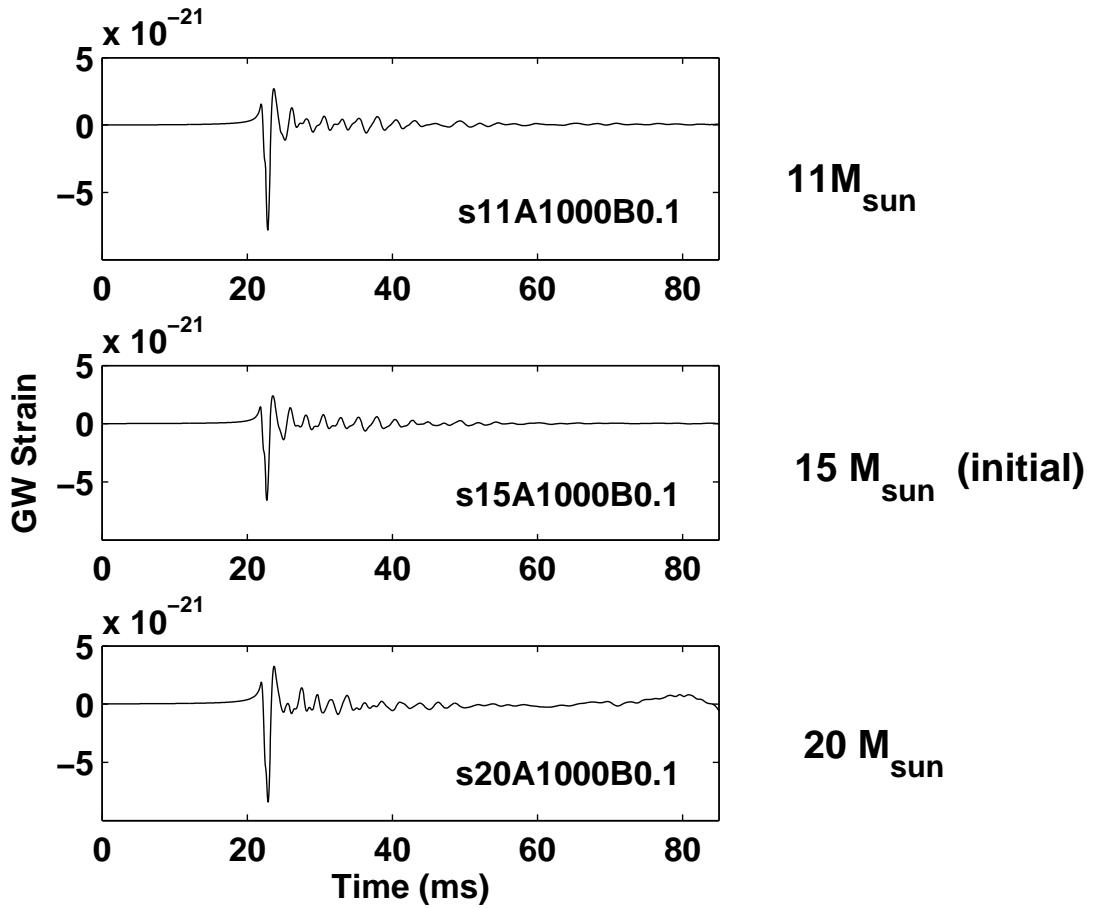


Fig. 8.— Waveforms from models that differ only by progenitor mass. The waveform corresponding to Ott et al. (2004) model s15A1000B0.1 was used as the initial signal in the detection simulations. The zero points of the time axes are chosen so that the minima of the waveforms occur at the same time for ease of comparison. The waveform amplitudes are scaled to correspond to core collapse events at 10 kpc.

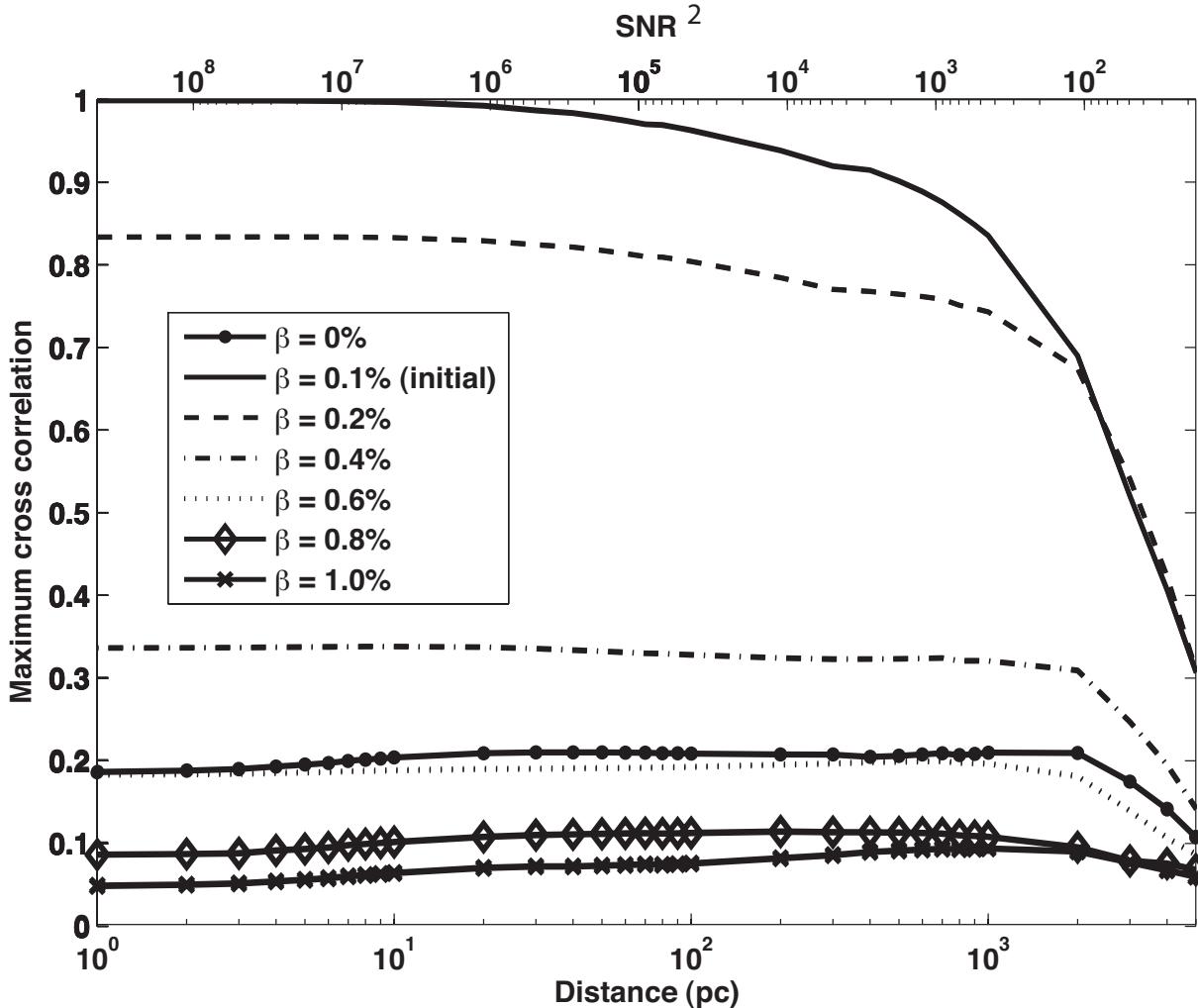


Fig. 9.— Maximum cross-correlation between reconstructed waveforms and waveforms associated with models that differ only by rotation parameter β , which is defined in equation (37), versus core collapse distance and SNR^2 . The SNR^2 shown is the average for the two detectors. The SNR^2 for the 4-km Hanford detector is 1.05 times that shown while the SNR^2 for the 4-km Livingston detector is 0.95 times that shown. The reconstructed waveform is inferred using maximum entropy from simulated detections that use a waveform from a model with a rotation parameter of $\beta = 0.1\%$ (Ott et al. [2004] model s15A1000B.0.1) as the initial signal waveform as well as detector responses and noise levels from the fourth science run (S4). The inferred waveform is most similar to that generated by the model with the same β for the simulations corresponding to core collapse events that occur less than 2-3 kpc away.

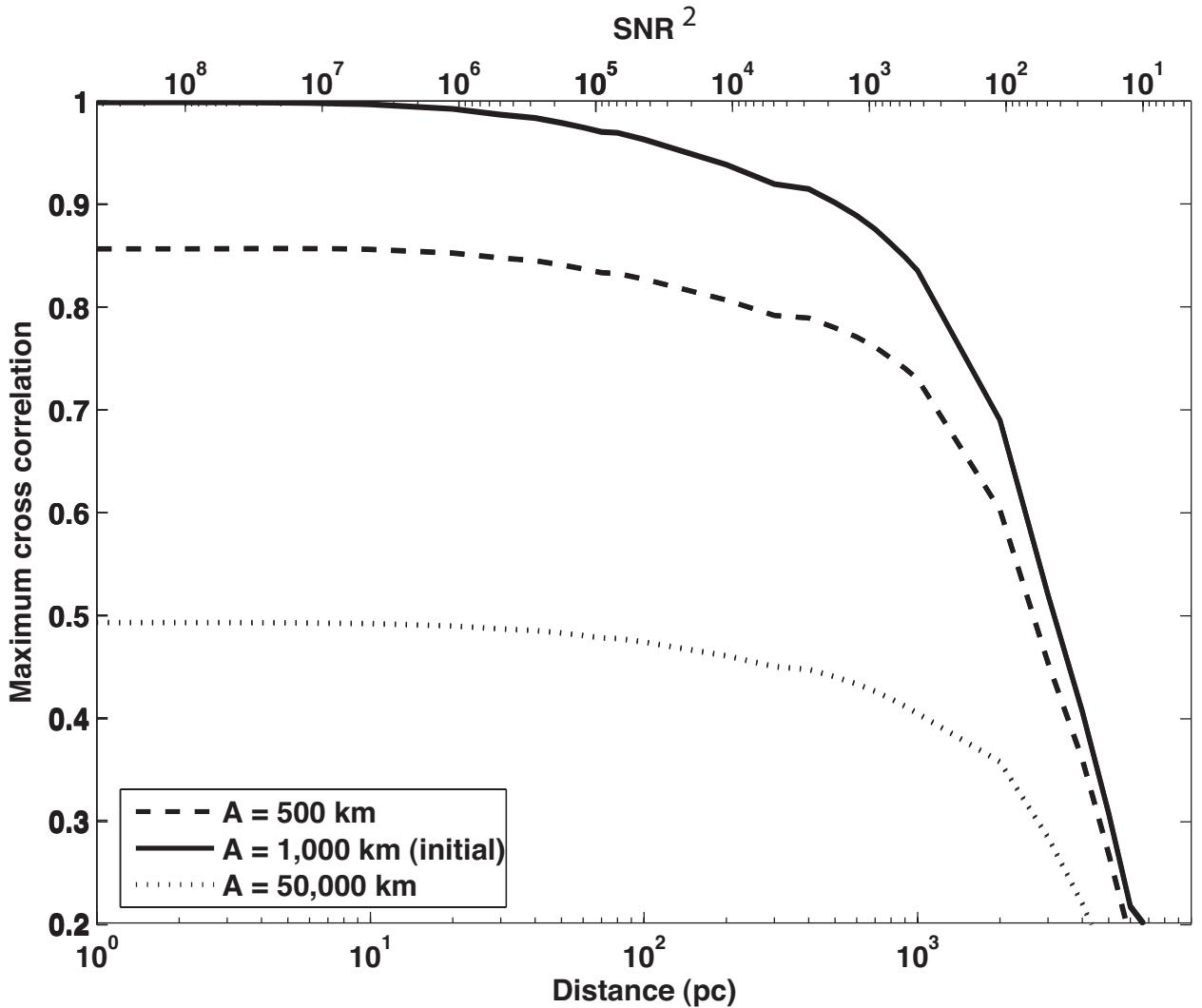


Fig. 10.— Maximum cross-correlation between reconstructed waveforms and waveforms associated with models that differ only by initial degree of differential rotation as parameterized by A , which is defined in equation (38), versus core collapse distance and SNR^2 . The SNR^2 shown is the average for the two detectors. The SNR^2 for the 4-km Hanford detector is 1.05 times that shown while the SNR^2 for the 4-km Livingston detector is 0.95 times that shown. The reconstructed waveform is inferred using maximum entropy from simulated detections that use a waveform from a model with a differential rotation parameter of $A = 1,000$ km (Ott et al. [2004] model s15A1000B0.1) as the initial signal waveform as well as detector responses and noise levels from the fourth science run (S4). The inferred waveform is the most similar to that generated by the model with the same initial degree of differential rotation.

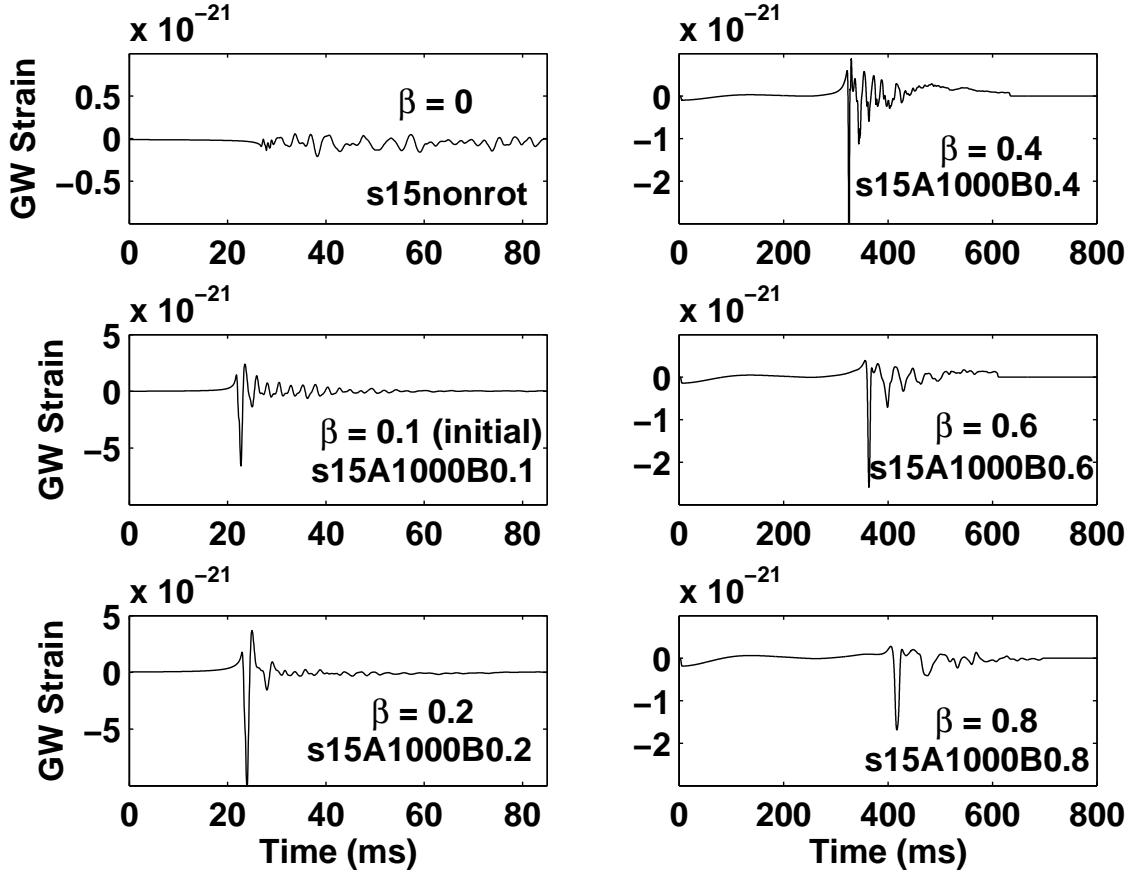


Fig. 11.— Waveforms from models that differ only by rotation parameter β , defined by equation (37). The waveforms for larger β ($\geq 0.4\%$) have significant amplitude over durations of hundreds of ms, while low β waveforms last only for tens of ms. The $\beta = 0\%$ waveform has very low amplitude as the non-rotating collapse is nearly spherically symmetric. The waveform corresponding to Ott et al. (2004) model s15A1000B0.1 was used as the initial signal in the detection simulations. The zero point of the time axes for the plots on the right is chosen so that the onset of significant gravitational wave amplitude occurs at roughly the same time while the plots on the left show the first 800 ms of the waveform. The waveform amplitudes are scaled to correspond to core collapse events at 10 kpc.

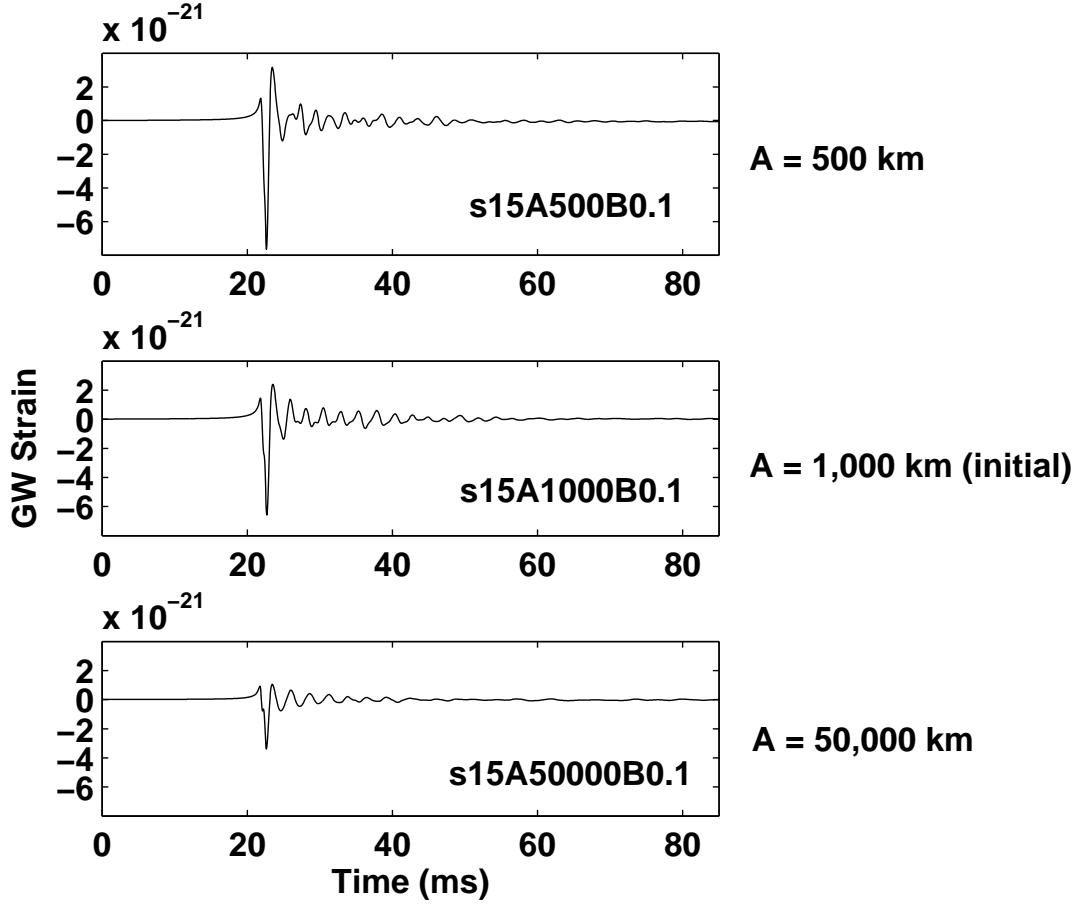


Fig. 12.— Waveforms from models that differ only by initial degree of differential rotation. The differential rotation parameter, A , is the distance at which the rotational velocity of the progenitor drops to half the rotational velocity at its center. As A decreases the differential rotation of the progenitor becomes more extreme and the amplitudes of the gravitational waves increase. The center plot shows the waveform corresponding to Ott et al. (2004) model s15A1000B0.1 which was used as the initial signal in the detection simulations. The zero points of the time axes are chosen so that the minima of the waveforms occur at the same time for ease of comparison. The waveform amplitudes are scaled to correspond to core collapse events at 10 kpc.